Uncertainty and Disagreement in Forecasting Inflation

Pfajfar, D.; Zakelj, B.

Publication date:
2012

Citation for published version (APA):
Pfajfar, D., & Zakelj, B. (2012). *Uncertainty and Disagreement in Forecasting Inflation: Evidence from the Laboratory (Revised version of EBC DP 2011-014)*. (EBC Discussion Paper; Vol. 2012-018). EBC.

Discussion paper

UNCERTAINTY AND DISAGREEMENT IN FORECASTING INFLATION: EVIDENCE FROM THE LABORATORY

By
Damjan Pfajfar, Blaž Žakelj

September 3, 2012

European Banking Center Discussion Paper
No. 2012-018

This is also a CentER Discussion Paper
No. 2012-072

This is a revised version of:
European Banking Center Discussion Paper
No. 2011-014

CentER Discussion Paper
No. 2011-053

ISSN 0924-7815

Uncertainty and Disagreement in Forecasting Inflation: Evidence from the Laboratory*

Damjan Pfajfar[†] Blaž Žakelj[‡]

September 3, 2012

Abstract

This paper compares the behavior of subjects' uncertainty in different monetary policy environments when forecasting inflation in the laboratory. We find that inflation targeting produces lower uncertainty and higher accuracy of interval forecasts than inflation forecast targeting. We also establish several stylized facts about the behavior of individual uncertainty, aggregate distribution of forecasts, and disagreement between individuals. We find that the average confidence interval is the measure that performs best in forecasting inflation uncertainty. Subjects correctly perceive the underlying inflation uncertainty in only 60% of cases and tend to report asymmetric confidence intervals, perceiving higher uncertainty with respect to inflation increases.

JEL: C91, C92, E37, D80

Keywords: Laboratory Experiments, Confidence Bounds, New Keynesian Model, Inflation Expectations

*We would like to thank Chryssi Giannitsarou, Cars Hommes, Seppo Honkapohja, Rosemarie Nagel, Karl Schlag and participants at the Bank of England, Bank of Canada, 2011 Florence Workshop on Behavioral and Exp. Economics, 2010 N.A. ESA Conference in Tucson, 2010 Money, Macro and Finance conference in Limassol, 2010 Computing in Economics and Finance conference in London, and 2010 EEA congress in Glasgow for their comments and suggestions. Blaž Žakelj gratefully acknowledges financial support of the Max Weber Fellowship. Damjan Pfajfar gratefully acknowledges funding from a Marie Curie project PIEF-GA-2009-254956 – EXPMAC. All remaining errors are ours.

†CentER, EBC, University of Tilburg, *Address:* Department of Economics, Faculty of Economics and Business Administration, P.O. Box 90153, NL-5000 LE Tilburg, Netherlands. *E-mail*: D.Pfajfar@uvt.nl. *Web*: https://sites.google.com/site/dpfajfar/.

‡Universitat Pompeu Fabra, Address: Universitat Pompeu Fabra, Ramon Trias Fargas 25-27, 08005 Barcelona, Spain. *Email*: blaz.zakelj@upf.edu.

1 Introduction

This paper discusses an experimental study on the expectation formation process and the associated uncertainty within a macroeconomic framework. The importance of inflation uncertainty has been recognized at least since Friedman's Nobel Lecture (Friedman, 1977). Friedman argued that higher rates of inflation are associated with higher inflation variability, which in turn causes a reduction in the efficiency of the price system and leads to a reduction in output due to institutional rigidities. Indeed, Levi and Makin (1980) and Mullineaux (1980) found empirical support for Friedman's conjecture. This represents a clear rationale for central banks to care about inflation uncertainty. Moreover, inflation targeting central banks, in particular, trust that inflation expectations of economic agents can be importantly shaped by their communication strategies. Both individual uncertainty and disagreement (interpersonal uncertainty) can be viewed as measures of the effectiveness of their communication strategies. In his speech about Federal Reserve communications, Mishkin (2008) stressed that the cost of inflation should be viewed both in terms of its level and of its uncertainty. As Giordani and Söderlind (2003) demonstrate this is particularly relevant when there is a regime switch (see also Evans and Wachtel, 1993). More generally, this is consistent with the standard New Keynesian dynamic stochastic general equilibrium (DSGE) model, in which a central bank should minimize the variation of inflation in order to maximize consumer welfare (see e.g., Woodford, 2003).[1]

In our experiment, subjects (undergraduate students) are introduced to a fictitious economy described by series of inflation, interest rates and the output gap. They are asked to forecast inflation and to provide 95% confidence intervals around their point forecasts. These forecasts are then fed into the simplified version of the New Keynesian model, which then generates realizations for inflation, the output gap and interest rates. These values are displayed to subjects and the process is iterated. This allows us to study both individual uncertainty about forecasts and disagreement on the point forecasts.[2]

[1] Recognizing the importance of different aspects of expectation distribution, Lorenzoni (2010) shows that monetary policy affects agents (with different pieces of information) differently, arguing that there is a tradeoff between aggregate and cross-sectional efficiency.

[2] Our companion paper (Pfajfar and Žakelj, 2011) focuses on the inflation expectation formation mechanism and its relation to monetary policy, i.e., how monetary policy should be designed in order to

In many respects it would be more desirable to study the responses of professional forecasters. They take economic decisions based on their forecasts and affect financial markets and the economy. However, there are clear advantages of using experimental data. We can: (i) observe how the forecasts interact with the fictitious model of the economy and different monetary policy regime, as we know the underlying model and the information set of survey respondents; (ii) analyze different policy regimes and different risk attitudes; (iii) have independent responses not affected by consensus opinion.

We focus on the relationship between monetary policy and inflation uncertainty and examine whether some environments are better than others at stabilizing inflation and minimize uncertainty. Two different monetary policy rules are evaluated: inflation targeting and inflation forecast targeting. For the latter we use three different specifications of the coefficient that describes the reaction of interest rates to deviations of inflation forecasts from the inflation target. We find that the monetary policy design significantly affects both the width and the accuracy of forecast intervals. In particular, the instrumental rule that reacts to current inflation reduces uncertainty and increases subjects' forecast accuracy compared to rules that react to expected inflation. Most of these differences can be attributed to the fact that certain monetary policy regimes result in lower variability of inflation than others. The contemporaneous rule (inflation targeting) produces a lower variability in actual inflation than inflation forecast treatments. Also treatments where the central bank reacts more strongly to deviations in inflation expectations from the inflation target result in lower inflation variability compared to the inflation forecasting regime with the baseline calibration. However, there are some treatment effects that go beyond this channel. Similar evidence is observed for different aggregate measures of uncertainty.

Different measures have been used in the survey data literature to proxy inflation uncertainty: the standard deviation of point forecasts, the average individual uncertainty, the interquartile range of the aggregate distribution of inflation forecasts (IQR), the average individual forecast error variance, and the variance of the aggregate distribution.

be robust to the potential presence of heterogeneous expectations.

Which of these representations of the distribution of inflation expectations is most relevant for the monetary authority?[3] Our analysis will mostly focus on the first three measures, since they are usually regarded as complementary in terms of informative value. For example, the first describes disagreement but says little about uncertainty, and the second captures uncertainty but disregards disagreement. Variance of the aggregate distribution of forecasts gives information about both uncertainty and disagreement. We find that average confidence intervals perform best in the forecasting exercise, although simple correlation analysis shows that the IQR is the measure that has the highest correlation with the variability of inflation.

When looking at individual responses we also find that forecasters usually tend to underestimate the underlying uncertainty when forecasting inflation, as only 60% of the results fall within the specified 95% intervals. That subjects tend to report narrower confidence intervals than that asked for is a well-known fact, labelled the "overconfidence effect." Several dynamic panel data regressions have been designed to identify the determinants of the measures discussed above. The results of an analysis of individual confidence intervals suggest that the width of the confidence interval is highly inertial and, interestingly, it increases only when inflation is below the target level. However, our results show little evidence of different degrees of uncertainty in different phases of the business cycle. Disagreement among subjects measured with the standard deviation of point forecasts increases when the average group forecast error increases and when inflation is below the target level. Absolute forecast error and inflation also affect individual uncertainty although disagreement is arguably less inertial. All the factors that significantly affect the specification of uncertainty and disagreement are by definition also important for the interquartile range. Indeed, inflation, the mean forecast error and the lagged interquartile range exert significant effects.

There are several reasons why asking subjects for symmetric intervals might be prefer-

[3] The forecasting ability of different measures has mostly been examined using the survey data of professional forecasters, see e.g., Zarnowitz and Lambros (1987) and Giordani and Söderlind (2003). See Engelberg, Manski, and Williams (2009) for some methodological issues involved. Zarnowitz and Lambros (1987), for example, show that there can be substantial differences between the variation in disagreement and the variation in uncertainty (for US survey of professional forecasters).

able to asking for potentially asymmetric intervals. Symmetric intervals are easier to handle in empirical analysis when the aim is to construct the aggregate distribution of expectations, because it can simply be assumed that an individual's distribution is normally distributed. Furthermore, there are no model based reasons why confidence intervals should not be symmetric, as the underlying model and the distribution of shocks do not exhibit any asymmetries. We have decided to perform treatments with a restriction to symmetric confidence intervals (we call them "Sym"), and treatments where we allow subjects to have potentially asymmetric intervals (we call them "Asym"). For the latter case, we find that only 12.5% of reported confidence intervals are symmetric. There is less asymmetry when there is an upward path of the output gap (expansion) and when inflation is below the target level.

Individual uncertainty and disagreement have not been extensively studied in economics using experimental data.[4] Forecasting uncertainty has attracted a lot of attention among psychologists, however, their approach has substantially different focus than ours.[5] Psychology literature usually limits its attention to independent event forecasts, while the present study concentrates on a series of (dependent) forecasts. This allows us to perform a time-series analysis of confidence bounds. We also provide subjects with other relevant information (besides the past history of prices) that might influence confidence. In this way we are able to examine whether confidence intervals are affected by the different designs of monetary policy.

This paper is organized as follows: Section 2 describes the model and the experimental design; in Section 3 we focus on the analysis of the individual responses while in Section 4 we analyze disagreement and the properties of aggregate distribution; Section 5 assesses the forecasting ability of different measures, while Section 6 concludes.

[4] Fehr and Tyran (2008) ask subjects to provide descriptive measures of their confidence level (but do not perform any analysis of them), while we ask subjects to provide numerical responses. Similarly, Bottazzi and Devetag (2005) ask subjects to provide 95% confidence intervals in an asset pricing experiment, with the aim (almost exclusively) of defining the average forecast but not of studying the behavior of uncertainty or disagreement.

[5] A common approach in this experimantal literature is to frame the question in the context of stock market forecasting exercises. For surveys, see Hoffrage (2004) or Lichtenstein, Fischhoff, and Phillips (1982) (see also e.g., Oskamp, 1965, Lawrence and O'Connor, 1992, Muradoglu and Onkal, 1994, Gilovich, Griffin, and Kahneman, 2002). These studies do not usually provide payment for the accuracy or the width of the confidence intervals, only for the accuracy of the point forecasts.

2 Experimental design

We design an experiment where subjects participate in a fictitious economy and are asked to provide inflation forecasts and a measure of uncertainty about their forecasts. The mean of the point forecasts is then used by the data generating process to calculate inflation, the interest rate, and the output gap. These variables are available to subjects before the next period forecast. Such so-called "learning to forecast" experiments have been conducted before within a simple macroeconomic setup (e.g., Williams, 1987; Marimon, Spear, and Sunder, 1993; Evans, Honkapohja, and Marimon, 2001; Arifovic and Sargent, 2003) and also within the asset pricing framework (see Hommes, Sonnemans, Tuinstra, and van de Velden, 2005 and Anufriev and Hommes, 2012). Closest to our framework, but with a different focus, are experiments by Adam (2007) and Assenza, Heemeijer, Hommes, and Massaro (2011). In this paper we have decided to focus on the reduced form of the New Keynesian (NK) model, where we can clearly observe forecasts and study their relationship with monetary policy.

Of course, there is a trade-off between using the model from "first principles" and employing a reduced form as in the learning to forecast literature. The former has the advantage of setting the objectives (payoff function) exactly in line with the microfoundations (see Noussair, Pfajfar, and Zsiros, 2011). In such approach, however, it is difficult to elicit forecasts since subjects act as producers and consumers and interact on the labor and final product markets and do not explicitly provide their forecasts.[6] We first present the model and then focus on the design.

The data generating process is a forward-looking sticky price NK monetary model with different monetary policy reaction functions.[7] The model consists of a forward-

[6]The argument is similar to that in papers by Marimon and Sunder, where the same tradeoff was first recognized. In this framework, thus, we do not assign subjects a particular role in the economy, rather they act as "professional" forecasters. One way to think about the relation between "professional forecasters" and consumers/firms is that these economic subjects employ professional forecasters to provide them with forecasts of inflation and associated uncertainty.

[7]The advantage of this small-scale NK model is that it reproduces relatively well several stylized facts about major economies and is the simplest model that is widely used for policy analysis. However, it requires forecasting two periods ahead. It would definitely be easier for participants to produce a one-period-ahead forecast (sometimes called "nowcasting"). The second drawback is that in standard NK models agents have to forecast both inflation and the output gap. We have decided to simplify this experiment by asking only for expectations of inflation.

looking Phillips curve (PC), an IS curve, and a monetary policy reaction function.[8] The information set at the time of forecasting consists of macro variables at time $t-1$, although the forecasts are made in period t for period $t+1$. Mathematically we denote this as $E_t(\pi_{t+1}|\mathcal{I}_{t-1})$, or simply $E_t\pi_{t+1}$. In the experiment we calculate it as $\frac{1}{K}\sum^k \pi^k_{t+1|t}$, where $\pi^k_{t+1|t}$ is subject k's point forecast of inflation (K is the total number of subjects in the economy). The IS curve is specified as follows:

$$y_t = -\varphi\left(i_t - E_t\pi_{t+1}\right) + y_{t-1} + g_t, \tag{1}$$

where i_t is the interest rate, π_t denotes inflation, y_t is the output gap, and g_t is an exogenous shock. The parameter φ is the intertemporal elasticity of substitution in demand. We calibrate it to 0.164.[9] It can be observed that we do not have expectations of the output gap in the specification. Instead, we have a lagged output gap.[10] Compared to purely forward-looking specifications, our model might display more persistence in the output gap. This is the most significant departure from the otherwise standard macroeconomic model. The supply side of the economy is summarized in the following PC:

$$\pi_t = \lambda y_t + \beta E_t\pi_{t+1} + u_t. \tag{2}$$

The longer prices are fixed on average, i.e., the smaller λ is, the less sensitive inflation is to the current output gap. The parameter β is the subjective discount rate. We set $\lambda = 0.3$ and $\beta = 0.99$. The shocks g_t and u_t are uncorrelated and unobservable to subjects and follow an AR(1) process: $g_t = \kappa g_{t-1} + \widetilde{g}_t$; $u_t = \nu u_{t-1} + \widetilde{u}_t$, where $0 < |\kappa|, |\nu| < 1$. $\widetilde{g}_t$ and $\widetilde{u}_t$ are independent white noise, $\widetilde{g}_t \backsim N\left(0,\sigma_g^2\right)$ and $\widetilde{u}_t \backsim N\left(0,\sigma_u^2\right)$. In the NK literature it is standard to assume AR(1) shocks. In particular, κ and ν are calibrated to 0.6, while their standard deviations are 0.08. The treatments are fully comparable as we

[8] Detailed derivations can be found in, e.g., textbooks such as Walsh (2003) or Woodford (2003).

[9] We implement McCallum and Nelson's (1999) calibration.

[10] In principle, one could argue that this specification of the IS equation corresponds to the case when subjects have naive expectations on the output gap or an extreme case of habit persistence is assumed. We were afraid that if we asked ubjects to forecast both inflation and the output gap this would represent a too difficult task for them (particularly given that they have to forecast for two periods ahead).

have exactly the same shocks in all treatments.[11]

Monetary policy reaction functions define four different types of treatments: three with (i) inflation forecast targeting (eq. 3) and different levels of γ, and one with (ii) inflation targeting (eq. 4):

$$i_t = \gamma\left(E_t \pi_{t+1} - \overline{\pi}\right) + \overline{\pi}, \quad (3)$$

$$i_t = \gamma\left(\pi_t - \overline{\pi}\right) + \overline{\pi}. \quad (4)$$

The central bank responds to deviations of forecasted inflation or of contemporaneous inflation from the target, $\overline{\pi}$. In order to have inflation in positive numbers for most of the periods, we set the inflation target to $\overline{\pi} = 3$. The aggressiveness of the central bank's response is determined by the reaction coefficient γ. A higher γ implies a stronger stabilizing effect of the Taylor-type monetary policy rule.

There are eight treatments in our experiment, depending on a monetary policy regime and a type of the confidence interval input. Treatments where subjects are asked to provide a symmetric confidence interval are denoted as Sym-p, where p represents one of the four monetary policy regimes. Subjects report the difference from their point forecast, which is roughly equivalent to 1.96 standard errors of their expectation, assuming it is represented by a normal distribution. Treatments where subjects are not restricted to symmetric confidence intervals are denoted as Asym-p, $p = 1, \ldots, 4$. In this set of treatments, subjects have to report the upper and the lower bound of their forecast together with their mean forecast, so that we do not restrict individuals to report symmetric confidence intervals (in both treatments we ask them to report 95% confidence intervals). Table 1 provides a summary.

Insert Table 1 about here

[11] g_t could be justified as government spending shocks or taste shocks and the standard interpretation of u_t is the technology shock. The empirical literature finds these shocks to be quite persistent (see e.g. Cooley and Prescott, 1995 or Ireland, 2004). The introduction of shocks as an exogenous unobservable component in the law of motion for observable variables is an important source of uncertainty in our experiment. It helps to avoid outcomes where all agents would implicitly coordinate on the forecasts identical to the inflation target and maintains the focus on the process of learning to forecast.

As said, three different levels of γ are tested in our experiment. We chose $\gamma = 1.5$ as our baseline specification in line with the majority of empirical findings and the initial proposal of Taylor in 1993. Our key interest is in seeing how subjects react to more ($\gamma = 4$) and less ($\gamma = 1.35$) aggressive interest rate policies and how these polices influence the uncertainty of their forecasts (see Section 3). A detailed discussion of the treatment selection regarding monetary policy and stability properties of the respective models can be found in our companion paper (Pfajfar and Žakelj, 2011), where we focus on questions regarding the performance of different monetary policy rules and their impact on expectations (point forecasts). However, let us point out that all treatments are determinate and E-stable under rational expectations. In our analysis we sometimes group all symmetric and asymmetric treatments together. We refer to treatments Sym-1, . . . , Sym-4 as Sym, and to treatments Asym-1, . . . , Asym-4 as Asym.

2.1 Experimental procedures

The experimental subjects participate in a simulated economy of 9 agents.[12] The participants were enlisted through a recruitment program for undergraduate students at the Universitat Pompeu Fabra and the University of Tilburg. The experiment consists of 24 independent groups of 9 subjects, 216 subjects in total. Each subject was randomly assigned to one group only. They earn on average around €15, depending on the treatments and individual performance. The participants receive detailed instructions (here attached in Appendix E) and a quiz questionnaire, and play 5 practice rounds before the start of the experiment to make sure they fully understand their task. The program is written in Z-Tree experimental software (Fischbacher, 2007).

The subjects are presented with a simple fictitious economy setup. As shown above, the economy is described by three macroeconomic variables: inflation, the output gap and the interest rate. The participants observe time series of these variables and their past forecasts, up to the period $t-1$. They do not observe the forecasts of other individuals or

[12] The common view among experimental economists is that it is not necessary to have many subjects in microfounded experiments. Most learning to forecast experiments are conducted with 5-6 subjects, e.g. Hommes, Sonnemans, Tuinstra, and van de Velden (2005), Adam (2007), Fehr and Tyran (2008).

their performance. 10 initial values are generated by the computer under the assumption of rational expectations. The subjects' task is to provide inflation forecasts for the period $t+1$ with a 95% confidence interval. Confidence intervals do not influence outcomes. The underlying model of the economy is qualitatively described to them. We explain the meaning and relevance of the main macroeconomic variables and inform them that their decisions have an impact on the output, inflation and interest rate realized at time t. This is the predominant strategy in learning to forecast experiments (see Duffy, 2008, and Hommes, 2011).[13] Each session consists of 70 periods.

After each period subjects receive information about the inflation realized in that period, their prediction of it, and the payoff they have gained. The payoff function is the sum of two convex components. The first component depends on their forecast errors, while the second depends on the width of their confidence interval.

$$
\begin{aligned}
W &= W_1 + W_2, \\
W_1 &= \max\left\{\frac{100}{1+f} - 20,\ 0\right\},\ W_2 = \max\left\{\frac{100x}{1+CI} - 20,\ 0\right\}, \\
x &= \begin{cases} 1 & \text{if} \quad CI \geq f \\ 0 & \text{if} \quad otherwise \end{cases},\quad f = \left|\pi_t - \pi^k_{t+1|t}\right|.
\end{aligned}
$$

The first component, W_1, depends on their forecast errors and is designed to encourage subjects to give accurate point forecasts. It gives subjects a payoff if their forecast errors, f, are smaller than 4. The second component, W_2, depends on the width of their confidence interval and is intended to motivate subjects to think about the variance of the actual inflation since it is more rewarding when it is lower. There is thus a trade-off between the width of this interval and its accuracy. A similar functional form of the

[13] In learning to forecast experiments it is not possible to achieve a REE (Rational Expectations Equilibrium) simply by introspection. This holds even if we provide subjects with the data generating process, as there exists uncertainty over how other participants forecast, so subjects have to engage in a number of trial and error exercises or in other words adaptive learning. It has been analytically proven in Marcet and Sargent (1989) and further formalized in a series of papers by Evans and Honkapohja (see their book: Evans and Honkapohja, 2001) that it is enough that agents observe all the relevant variables in the economy (as in our case, where they are specifically instructed that all of them might be relevant) and update their forecasts according to the adaptive learning algorithm (their errors) for them to end up in the REE. This has been acknowledged also in Duffy (2008) and Hommes (2011).

payoff function is used in Adam (2007). The CI is either equal to the subjects' point estimate of their confidence interval or half of the difference between the upper and lower bounds. The subjects receive a reward if their confidence intervals, CI, are not larger than ± 4 percentage points,[14] conditional on the fact that actual inflation falls in the given interval: $CI \geq \left|\pi_t - \pi^k_{t+1|t}\right|$. With this setup we restrict to positive payoffs.

Prior to running the experiments we investigated alternative designs of the payoff function W_2. Desirable payoff functions should maximize the forecaster's utility at confidence intervals giving some fixed proportion of the forecast uncertainty distribution. Confidence intervals are more informative when they encompass a relatively high proportion of forecaster's uncertainty distribution. We therefore decided to aim for the interval that includes (roughly) 95% of the inflation forecast uncertainty distribution. It is difficult, however, to specify and calibrate a payoff function that is maximized at the required proportion of the inflation forecast uncertainty distribution and is at the same time easily understandable to the subjects. Furthermore, subjects may use alternative forecasting models for point predictions with different distribution variances. As shocks are not directly observable in our experiment, the relevant benchmark to study the properties of the payoff function is the restrictive perception equilibrium, which is of the same form as the rational expectation equilibrium, except that agents do not observe the shocks. Under this expectation formation mechanism the payoff function that has been chosen is maximized when providing $89 - 94\%$ confidence intervals, depending on the policy rule implemented. In Appendix D, our payoff function W_2 is studied in detail.

3 Individual uncertainty

While the distribution of means across subjects captures interpersonal variation, confidence bounds help us to approximate individual uncertainty of future inflation. Zarnowitz and Lambros (1987) show that there can be substantial differences between the variation in disagreement and the variation in average uncertainty in US Survey of Professional Forecasters (SPF). Therefore, both might not be appropriate measures for forecasting

[14] This constraint was almost never binding (see Figure 1). In only 0.98% cases this was violated.

inflation uncertainty. The current section concentrates on individual uncertainty, while the next section investigates the aggregate distribution of forecasts and disagreement. In both sections a clear emphasis is posed on the behavior of these measures across different designs of monetary policy.

Insert Figure 1 about here

Figure 1 displays the distribution of all confidence interval forecasts. The range of responses for confidence intervals is between 0 and 8.3, although it should be noted that responses larger than 4 do not result in any payoff (0.98% of all cases). The average symmetrical confidence interval is 0.61, with an average standard deviation of 0.28. Allowing for asymmetrical confidence bounds across all policy regimes we obtain an average lower part of the confidence interval of 0.37 with an average standard deviation of 0.19, while the average upper part of the confidence interval equals 0.41 with an average standard deviation of 0.28. There are considerable differences across treatments as the lowest symmetrical (asymmetrical lower, upper) average interval in treatments Sym (treatments Asym) is 0.41 (0.24, 0.27) and the highest is 0.91 (0.47, 0.53). Evidence of rounding is present in responses 0.5, 1, 1.5, 2, and 3 as they have significantly higher frequencies than other responses. Overall, 13% of responses are integers, while the majority are to one decimal point accuracy, 77%. The remaining responses are to 2 decimal point accuracy.

Insert Table 2 about here

The average confidence intervals in all treatments are listed in Table 2, while a per-group summary is presented in Table 3. In general, confidence intervals are narrower in treatments Asym than in treatments Sym at 1% significance using nonparametric tests (Wilcoxon/Mann-Whitney). In Section 3.1 we show that treatments Sym and Asym also differ in the forecast accuracy of subjects' interval predictions. The factors that determine the differences in confidence intervals are discussed in Section 3.2.[15]

[15] Our results might not be directly comparable to those based on surveys. Probabilistic forecasts in surveys are usually collected in terms of histograms where intervals are predefined and fixed for all participants. Another difference between our experiment and surveys concerns the risk attitude and the horizon of the forecasts. With professional forecasters it can be claimed that their probability and point forecasts are correlated. In our experiment, subjects could neither exchange information about each other's expectations, nor is the average aggregate prediction directly observable.

3.1 Monetary policy, individual uncertainty, and forecasting accuracy

We have the opportunity to compare the results with the underlying uncertainty that we have embedded in our set-up. In Appendix D we calculate, for all policy rules, confidence intervals that maximize the payoffs under rational expectations where agents observe the shocks, and restrictive perception equilibrium, where shocks are not observed which is consistent with our experimental design. Of course, as soon as one subject departs from rationality, the confidence interval of a rational agent should immediately become larger as she has to account for the uncertainty of other subjects' expectations. Under rational expectations in treatments with policy regimes $p = 1, \ldots, 3$ higher γ leads to lower uncertainty, while when we compare treatments with policy regimes $p = 4$ and $p = 1$, inflation targeting and inflation forecast targeting with $\gamma = 1.5$, the former should produce less uncertainty. These properties have testable implications.

Let us first focus on the conjecture that uncertainty should be lower when the central bank is pursuing inflation targeting compared to inflation forecast targeting. We indeed find that the average confidence interval is narrower in policy regime 4 compared to the regime 1. This difference is statistically significant with standard parametric (t-test) and nonparametric tests (Wilcoxon/Mann-Whitney). If we compare treatments with regime $p = 4$ separately to all the other regimes, we observe that while it is significantly narrower than in treatments with policy regimes 1 and 2 (i.e., $\gamma = 1.5$ and $\gamma = 1.35$), the confidence interval is wider than in treatments with policy regime 3 ($\gamma = 4$).

The second testable implication is that in treatments with inflation forecast targeting ($p = 1, 2, 3$) higher γ leads to lower uncertainty; this is also strongly confirmed by our experimental data. Using Jonckheere-Terpstra test we find that there is descending width of confidence intervals when we increase γ. Thus, using both parametric and non-parametric tests we can conclude that monetary policy significantly affects the width of the confidence interval. In specific, inflation targeting results in a narrower confidence interval than inflation forecast targeting and that in the case of inflation forecast targeting, the width of the confidence interval also depends on how strongly the monetary policy is reacting to deviations of inflation from its target. One possible explanation is that the confidence intervals depend on the variation of inflation. The variation of inflation is lower in inflation targeting treatments than inflation forecast targeting treatments with the same γ. Also treatments with policy regime 3 produce lower variability than treatments with policy regimes 1 and 2.[16] We further study this in Section 3.2.

Insert Table 3 about here

It is interesting to see how accurate experimental subjects are in determining the confidence bounds. Thaler (2000) claims that when people are asked "for their 90% confidence limits ... the correct answers will lie within the limits less than 70% of the time" (p. 133).[17] Our results confirm the overconfidence effect in an even stronger manner than survey data results. Only 60.5% of the times do subjects manage to set confidence bounds that include the actual inflation in the next period when asking for 95% confidence intervals.[18] This proportion is higher in treatments Sym (64.3%) than in treatments Asym (52.8%). It is interesting to note that the actual inflation is lower than their confidence intervals in 19% of cases while it is higher in 20.5%. If we compare this among treatments we find that in treatments Sym (Asym) actual inflation is lower than their confidence intervals in 17.1% (22.9%) of cases while it is higher in 18.5% (24.4%).[19]

The accuracy of confidence intervals also differs across different monetary policies. We find that in treatments with policy regime 3 (inflation forecast targeting, $\gamma = 4$) and 4 (inflation targeting), subjects are more accurate (62.9% and 69.4% accuracy respectively) than in the benchmark treatments with policy regime 1, inflation forecast targeting with $\gamma = 1.5$ (51.7% accuracy). The differences are significantly at a 10% level with the Wilcoxon/Mann-Whitney test. Thus monetary policy does not have only significant effects on the width of the interval, but also on the predictive accuracy.

[16] Table A3 reproduces these results from Pfajfar and Žakelj (2011) in the Appendix A.

[17] Giordani and Söderlind (2003) for the US SPF obtain a very similar result (72%). See also Giordani and Söderlind (2006) and Diebold, Hahn, and Tay (1999).

[18] As mentioned before, this is incentive compatible with $89-94\%$ confidence intervals under restricted perception equilibrium, depending on monetary policy regime.

[19] As mentioned in the introduction, this overconfidence effect has attracted a lot of attention in the psychology literature. Some studies even document that the success rate of these forecasts is less than 50% when people are asked for $90-99\%$ confidence intervals (e.g. Lichtenstein, Fischhoff, and Phillips, 1982).

As confidence intervals forecast the distribution of the expected forecast errors we can actually dig deeper and analyze each individual separately. We find that only 11.1% of the subjects on average overestimate risk in treatments Sym and 2.8% (1.4%) of the subjects in treatments Asym for the lower (upper) bound. Closer inspection allows us to conclude that on average only about 9.0% of the subjects in treatments Sym and 1.4% (8.4%) of the subjects in treatments Asym for the lower (upper) bound report confidence bounds that are not significantly different from 95% confidence intervals based on actual forecast errors. The rest of the subjects on average forecast confidence bounds that are significantly lower than the actual forecast errors.[20]

3.2 Determinants of individual uncertainty

Below we analyze the determinants of confidence bounds using panel data. All the regressions below are estimated using the system GMM estimator of Blundell and Bond (1998) for dynamic panel data. They are replicated for the whole sample (*all*), treatments Sym (*treat.Sym*), and separately for the part of the interval below the point forecast (*treat.Asym-L*) and above the point forecast (*treat.Asym-U*) in treatments Asym. In order to transform the asymmetric confidence intervals into a measure comparable to the symmetric ones, we compute the average of the upper and lower interval.

We begin by detailing the relationship between the confidence interval and the standard deviation of inflation. We estimate the following regression:

$$sip^k_{t+1|t} = \alpha + \beta sip^k_{t|t-1} + \gamma sd^j_{t-1} + u^{em}_t, \tag{5}$$

where individual k's current perceived uncertainty in period t, is measured by her confidence interval, $sip^k_{t+1|t}$ (for treatments Asym it is (Upper bound – Lower bound)/2). sd^j_{t-1} is the standard deviation of inflation up to period $t-1$ for group j. The results are reported in Table 4.

[20] Per-group statistics are reported in Table A1 in Appendix A, while in Table A2 the frequencies of forecast errors depending on the inflation cycle can be found. In the Appendix B.1 we also study the determinats of the likelihood that actual inflation falls within the specified bounds.

Insert Table 4 about here

We find that confidence intervals are highly inertial and that higher standard deviation of inflation leads to wider confidence intervals, although with a smaller effect in treatments Sym.[21] Table B5 in Appendix B demonstrates that if we include monetary policy dummies in regression (5) we find that the dummy variable for treatments with monetary policy 2 is significant. One reason behind this is that uncertainty is related to the variability of inflation, which in turn depends on γ and more generally on the monetary policy (see Pfajfar and Žakelj, 2011).[22] However, there also exist other monetary policy effects as can be observed when controlling for a standard deviation of inflation. This means that the relationship between monetary policy and uncertainty is not limited just to the channel via the standard deviation of inflation.

Several studies have established that uncertainty is countercyclical. Bloom (2009) builds a theoretical model where uncertainty shocks play a key role in business cycle fluctuations. Also Van Nieuwerburgh and Veldkamp (2006) propose an endogenous information model where macroeconomic uncertainty and dispersion in beliefs is greater during recessions. Although our model does not exhibit such features, i.e., is perfectly symmetric, some subjects that do not use all available information to forecast inflation and to report uncertainty might perceive asymmetric uncertainty over the business cycle. We estimate equation (6), where we control for the path of the output gap. In addition, specification (6) also allows for the possibility that subjects change their interval forecasts on the basis of their last point forecast errors:

$$\begin{aligned} sip^k_{t+1|t} &= \alpha + \beta sip^k_{t|t-1} + \gamma D_1 y_{t-1} + \delta D_2 y_{t-1} + \epsilon D_3 y_{t-1} \\ &\quad + \zeta i_{t-1} + \eta D_L \left|\pi_{t-1}\right| + \theta D_H \left|\pi_{t-1}\right| + \phi \left|r^k_{t-1}\right| + \vartheta T2 + \iota T3 + \kappa T4 + u^{em}_t, \end{aligned} \tag{6}$$

[21] Inertia of confidence intevals has been previously documented in the survey data literature by, e.g., Bruine de Bruin, Manski, Topa, and van der Klaauw (2011), and Giordani and Söderlind (2003). the former also find a positive correlation between the self-reported range of responses and the underlying uncertainty in US consumer data.

[22] Due to the presence of heterogeneous expectations this relationship is not monotonic. It is found that the relationship between γ and the variance of inflation is under certain expectation formation mechanisms U-shaped (see Pfajfar and Žakelj, 2011 for further details).

where y_t is the output gap, π_t is actual inflation, i_t is the interest rate in group j.[23] $D_1, \ldots, D_3$ are dummy variables. D_1 equals 1 when $y_{t-1} > 0.1$ and $\Delta y_{t-1} > 0$ and is 0 otherwise; D_2 equals 1 when $y_{t-1} < 0.1$ and $\Delta y_{t-1} < 0$ and is 0 otherwise; D_3 equals 1 when $D_1 = 0$ and $D_2 = 0$ jointly and is 0 otherwise. $T2$, $T3$ and $T4$ are dummies for treatments with policy regimes $1, \ldots, 3$ respectively.[24] D_L equals 1 when inflation is below the target and 0 otherwise, while D_H equals 1 when inflation is above its target and 0 otherwise.

Insert Table 5 about here

Results from eq. (6) are reported in Table 5, which shows that different monetary policy rules have an effect on the width of the confidence interval as already established above. The confidence intervals are wider for example in treatments with policy rule $p = 2$ compared to the treatments with other policy rules. Friedman (1968) points out that there is a positive link between inflation and inflation uncertainty. While, for example, Liu and Lahiri (2006) and D'Amico and Orphanides (2008) find empirical support for this conjecture in the survey data literature, we cannot confirm it in our experiment.[25] Regressing equation (6) with inflation (π_{t-1}) instead of $D_L\,|\pi_{t-1}|$ and $D_H\,|\pi_{t-1}|$ would result in inflation having a negative impact on the width of the confidence interval. The empirical studies that find a positive correlation between inflation and uncertainty are based on the US economy where, especially in the 70s, there was mostly an upward risk for inflation. In our experiment, inflation fluctuates around the inflation target, so decreases in inflation below the inflation target also increase uncertainty. With specification (6) we concentrate on the absolute deviations of inflation from the inflation target, while controlling for high and low inflation levels. We indeed observe that downside risk has an even more important impact on the uncertainty than the upside risk. Moreover, when

[23] To ease the notation we omit superscripts j for these variables.

[24] Monetary policy dummies are included only in regression *all* as in the other specifications due to too few observations within one treatment we would have to abolish the clustering of standard errors if we were to include policy dummies.

[25] Rich and Tracy (2010) and García and Manzanares (2007), among others, study the relationship between confidence intervals and inflation forecasts using survey data. They find mixed evidence of the existence of the relationship. We do not find evidence in favor of this relationship (these results are available upon request).

inflation is above the target inflation only the upper part of the confidence interval will be widened, whereas when it is below the target inflation both sides of the confidence interval will be widened.

Interest rates are positively related to the individual confidence intervals in the regressions above, although their effects are not significant. In Table 5 we also demonstrate that confidence intervals positively depend on the last observed absolute forecast error. Additionally, we confirm the asymmetries between the upper and lower confidence bound demonstrated in Table 2. We can argue that the upper bound is more sensitive to the stage of the business cycle than the lower bound.[26]

3.3 Which factors affect the choice of (a)symmetric confidence intervals?

We have already found several asymmetries between the formation of the upper and the lower confidence bounds in the previous section. We now turn our attention to the choice of asymmetric confidence interval by using data from treatments Asym. Let us first analyze the proportion of subjects that systematically choose either a wider interval above the point forecast as compared to the one below point forecast or vice-versa. It is clear from Table 6 that when subjects are given the option to choose an asymmetric confidence interval they often do so, especially in treatments Asym-1 and Asym-2. Moreover, among more than 40% of the subjects who systematically choose asymmetric intervals, fewer than 6% perceive higher uncertainty on the left-hand side of their point forecast. We can also observe that the proportion of subjects choosing symmetric intervals is the highest in treatment Asym-4 (inflation targeting). Table 6 shows that the behavior of subjects in the inflation targeting treatments is more in line with theory than in the treatments with

[26] In the Appendix B.2 we also study subjects' responses to inflation falling outside the confidence interval. The results shown in Table B3 suggest that subjects increased their confidence intervals after the last observed inflation is outside the interval. This holds for both "undershooting" and "overshooting." In the case of "overshooting." a negative coefficient δ implies that confidence intervals are widened after an "error." Confidence intervals do not change when inflation falls within the interval in the previous period. Confidence intervals in treatments Asym exhibit less inertia, especially at the upper bound, compared to treatments Sym. Moreover, the interval above the point forecast widens with both overshooting and undershooting while the interval below is more stable and responds only to undershooting. This also represents one potential source of observed asymmetries.

inflation forecast targeting.

Insert Table 6 about here

Now we turn our attention to the factors that determine the probability of an asymmetric interval. We first define $D_7 = 1$ if the upper interval has exactly the same width as the lower one and 0 otherwise. There are only about 12.5% of these cases. We observe, that 84% of the subjects gave their responses with one or two decimal point accuracy. It is therefore reasonable to define symmetry as $|C_L - C_U| \leq 0.1$; in this case we set $D_8 = 1$. According to this definition 47.2% of our responses in treatments Asym are approximately symmetric. We estimate the following regressions:

$$\begin{aligned} D_z &= \alpha + \beta sip^k_{t|t-1} + \gamma D_1 y_{t-1} + \delta D_2 y_{t-1} + \epsilon D_3 y_{t-1} + \zeta i_{t-1} \\ &\quad + \eta D_L \left|\pi_{t-1}\right| + \theta D_H \left|\pi_{t-1}\right| + \phi sd^j_{t-1} + u^{em}_t; \quad z \in \{7, 8\}. \end{aligned} \tag{7}$$

The results for the logit fixed effects estimator are reported in the first two columns of Table 7.[27] While the above regressions inform us about the likelihood that subjects choose symmetric intervals, they are not suitable for measuring the magnitude of the asymmetry of the individual forecast distributions or their direction. For that purpose it is convenient to introduce a new variable, skewness, similar to that used in Du and Budescu (2007). We define the proxy for skewness, skw^k_t, by subtracting the point forecast from the midpoint of the confidence interval. If skw^k_t is smaller (greater) than 0, then the interval is left (right) skewed, and the confidence interval below the point forecast is wider (narrower) than the one above. If $skw^k_t = 0$ then the interval is symmetric. The factors affecting skewness are analyzed on the right-hand side of Table 7 using the Blundell-Bond system GMM estimator.

$$\begin{aligned} skw^k_t &= \alpha + \eta skw^k_{t-1} + \beta sip^k_{t|t-1} + \gamma D_1 y_{t-1} + \delta D_2 y_{t-1} + \epsilon D_3 y_{t-1} \\ &\quad + \zeta i_{t-1} + \eta D_L \left|\pi_{t-1}\right| + \theta D_H \left|\pi_{t-1}\right| + \phi sd^j_{t-1} + u^{em}_t. \end{aligned} \tag{8}$$

[27] Poisson and logit random effects estimations can be found in Tables B7 and B8 in Appendix B.3. Alternatively, we also tried different definitions of symmetry. $D_9 = 1$ if $0.9 \leq \left|\frac{ConfIntH_{n-1}}{ConfIntL_{n-1}}\right| \leq 1.1$. The results can be found in Tables B7-9 in Appendix B.

Insert Table 7 about here

Regressions for D_7 and D_8 demonstrate that some indicators of the cycle are significant. In particular, for D_7 when the output gap is negative and downward sloping to observe symmetric intervals it is less likely, while for D_8 observing symmetrical intervals in the opposite stage of the business cycle is more likely. For both regressions, the interest rate has a significantly positive impact and absolute inflation above the target a significantly negative impact, i.e., there is less symmetry when inflation is low.

The skewness measure, on the other hand, also gives us an indication of the direction of the asymmetry. We find that this measure is inertial and tends to decrease (left skewness) when the previous confidence interval was larger. The measure also varies across the business cycles: it is lower when $D_3 = 1$. Du and Budescu (2007) find a negative relationship between the standard deviation of inflation and the skewness of confidence distribution, while we find this relationship only for the case of D_7.

4 Disagreement and aggregate expectation distribution

We first analyze the features of the standard deviation of point forecasts. Second, we take account of individual uncertainty as well. We define the probability density functions of individual distributions, add them up, and analyze the features of aggregate distribution.

4.1 Disagreement

Variance of point forecasts is a "natural" measure of disagreement. It is often used in the empirical literature since the data on point forecasts are more frequently available than the data on individual distributions. It is studied in the survey data literature, for example, in Zarnowitz and Lambros (1987), Giordani and Söderlind (2003), Patton and Timmermann (2008), and Rich and Tracy (2010). We investigate the relation of the standard deviation of point forecasts to the phases of the economic cycle, interest rate, inflation and the mean forecast error:

$$
\begin{aligned}
sdv^{j}_{t+1|t} &= \alpha + \beta sdv^{j}_{t|t-1} + \gamma D_1 y_{t-1} + \delta D_2 y_{t-1} + \epsilon D_3 y_{t-1} \\
&\quad + \zeta i_{t-1} + \eta D_L \left|\pi_{t-1}\right| + \theta D_H \left|\pi_{t-1}\right| + \phi m r^{j}_{t-1} + u^{em}_{t},
\end{aligned} \tag{9}
$$

where $sdv^{j}_{t+1|t}$ is a cross-sectional standard deviation of point forecasts in group j at period t, while the mean absolute forecast error in group j at period $t-1$ is mr^{j}_{t-1}.

The regressions based on (9) are displayed on the left-hand side of Table 8. The standard deviation of point forecasts exhibits sensitivity to inflation, mean absolute forecast error and to some degree business cycles. However it tends to be less sensitive to these variables in treatments with asymmetric confidence intervals, where only inertia and sensitivity to the business cycle play an important role. Disagreement increases when the output gap is below the steady state and falling. We observe higher disagreement when absolute inflation is below the target.[28]

Monetary policy produces significantly different median standard deviation of point forecasts (sdv) across treatments. Wilcoxon\Mann-Whitney test suggest that among inflation forecast targeting treatments sdv is the highest in treatments with policy regime 2 ($\gamma = 1.35$), followed by treatments with policy regime 1 ($\gamma = 1.5$), and the lowest in treatments with policy regime 3 ($\gamma = 4$). Furthermore, inflation targeting treatments (policy regime 4) produces lower sdv than comparable inflation forecast targeting treatment (policy regime 1). There are some treatment differences regarding the determination of the standard deviation of point forecasts (sdv). In particular, treatments with policy rule 3 seem to produce lower sdv compared to treatments with policy rule 1, therefore in the environment where monetary policy more strongly fights inflation the perceived uncertainty is lower. However, we are not able to introduce policy rule dummies to the regressions

[28]Cukierman and Wachtel (1979), Rich and Tracy (2010), D'Amico and Orphanides (2008), Capistrán and Timmermann (2009), and Mankiw, Reis, and Wolfers (2004) find that there is a positive relationship between inflation and disagreement. Our results conversely point out that low (below target) inflation can also generate higher uncertainty. Carroll (2003) and Mankiw, Reis, and Wolfers (2004) explain this result with sticky information model. Capistrán and Timmermann (2009) and Mankiw, Reis, and Wolfers (2004) also point out that the disagreement increaseses with the variance of current inflation. Patton and Timmermann (2008) demonstrate for US SPF that disagreement is countercyclical, being higher in recessions.

for the sdv and IQR as then we would not be able to compute clustered standard errors across treatments with different policy rules. These results are from estimations of eq. (9) with policy rule dummies using robust standard errors.[29]

Insert Table 8 about here

4.2 Dispersion of aggregate distribution

Several central banks have started to put the data on the forecast distribution of economic variables (including inflation expectations) on the agenda for policy meetings (see e.g., recent FOMC minutes). This is partly a product of advances in Bayesian estimation methods for monetary models and also of the adoption of new communication strategies by many central banks (see e.g., ECB Monthly Bulletin, 2012). One option is to aggregate individual distributions and analyze them, rather than calculate averages from the individual moments. We derive the distribution from the asymmetric confidence bounds by using a triangles approach similar to Engelberg, Manski, and Williams (2009). The mode is set to be equal to the point forecast, while 95% of the derived triangular distribution is set to be between the lower and the upper confidence bound. In this way we generate probability density functions for each forecast by an individual. The distributions are then aggregated (cross-sectionally) across the individuals in a group.

We choose the interquartile range (IQR), a range between the 25^{th} and 75^{th} percentile, as an appropriate measure as it is less sensitive to small variations in the tails of the estimated density compared to the cross-sectional standard deviation of the aggregate distribution.[30] Nevertheless, it is useful to show that the variance of aggregate distribution is related to the two measures that we study above. Wallis (2004, 2005) show explicitly that the variance of the aggregate distribution can be decomposed into the average individual uncertainty and disagreement of point forecasts. Comparing the

[29] Only Capistrán and Ramos-Francia (2010) have studies the relationship between monetary policy design and disagreement. In their sample of 24 countries among which 14 inflation targeters, they found that countries that adopted inflation targeting disagreement in long-run inflation expectation is smaller.

[30] Giordani and Söderlind (2003) use a similar measure to ours. In the literature other measures have also been proposed. Boero, Smith, and Wallis (2008) use the standard deviation of the aggregate distribution, while Batchelor and Dua (1996) suggest root mean subjective variance.

IQR across monetary policy regimes we find analogous evidence as in the case for *sdv*: among treatments with inflation forecast targeting the highest IQR is in treatments with policy regime 2, followed by treatments with policy regime 1, and the lowest in treatments with policy regime 3 ($\gamma = 4$). Also treatment with inflation targeting (policy regime 4) produces lower IQR than treatments with policy regime 1. To discover the properties of the aggregate distribution, we run the following regression:

$$\begin{aligned} IQR_t^j &= \alpha + \zeta IQR_{t-1}^j + \beta D_1 y_{t-1} + \gamma D_2 y_{t-1} + \delta D_3 y_{t-1} \\ &\quad + \epsilon i_{t-1} + \eta D_L \left|\pi_{t-1}\right| + \theta D_H \left|\pi_{t-1}\right| + \eta m r_{t-1}^j + u_t^{em}, \end{aligned} \tag{10}$$

where $IQR^j = Q_3 - Q_1$ is the interquartile range, y_t is the output gap, i_t is the interest rate, and $D_1, \ldots, D_3$ are dummy variables as identified above.

Equation (10) considers the sources of divergences in expectations, such as the output gap, the interest rate and the previous value of the interquartile range. As above, we introduce a dummy variable for each of the phases of the cycle. Several studies in the survey data literature observe considerable inertia in the disagreement of expectations (see e.g., Giordani and Söderlind, 2003). We therefore also include the previous period interquartile range among the independent variables and find them highly significant. The results on the right-hand side of Table 8 show that there is some influence of the cyclical phase and inflation on the interquartile range. For a negative and decreasing output gap there is more disagreement.[31] We observe that the interquartile range is positively correlated with the absolute level of inflation when inflation is below the target level. In treatments Sym, the mean absolute forecast error also significantly affects the IQR. It is worth noting that regressions for the treatments with symmetric and asymmetric confidence intervals show very similar results. Regression results yield no significant differences between the different monetary policy rules employed, when we control for the above regressors.

[31] This is similar to the results in survey data, where it is common to observe countercyclical behavior of variance of inflation expectations. Pfajfar and Santoro (2010) also study the kurtosis and skewness of the distribution of forecasts and find that both exhibit procyclical behavior.

5 Forecasting ability of different measures of inflation uncertainty and disagreement

Policymakers and researchers are interested in inflation uncertainty and in obtaining proxies for it. Therefore, the question that needs to be addressed is which measure or a combination of them best forecasts inflation uncertainty.[32] As we can observe in Table B9, the highest correlation is between the interquartile range (IQR) and the standard deviation of inflation (sd). It reaches almost 0.9, while somehow surprisingly disagreement is a slightly better proxy of inflation uncertainty than the average perceived uncertainty of subjects.[33] In order to further assess the forecasting performance of these measures we estimate the following regression:

$$\begin{aligned} sd_t^j &= \alpha + \beta sd_{t-1}^j + \gamma asip_{t-1}^j + \epsilon sdv_{t-1}^j + \delta IQR_{t-1}^j \\ &\quad + \zeta i_{t-1} + \eta \pi_{t-1} + \phi y_{t-1} + u_t^{em}, \end{aligned} \tag{11}$$

where $asip_{t-1}^j$ is the average confidence interval in period $t-1$ for group j. Table 9 reports the results. We estimate three different specifications, which are a subset of the above equation. In variant (a) we include all three measures, while in variant (b) we include only measures of individual uncertainty and disagreement. Variant (c) embeds only the IQR as it is a measure of both individual uncertainty and disagreement and, as pointed out above, it is the measure that has the highest correlation with the standard deviation of inflation.

Insert Table 9 about here

The regressions confirm that the average individual uncertainty and the standard deviation of point forecasts have a positive effect on inflation variance. It comes as a surprise however that the interquartile range has a marginally significant negative effect.

[32] Several studies have demonstrated that inflation uncertainty is useful for macroeconomic forecasting, see e.g., Ang, Bekaert, and Wei (2007), Söderlind (2011), and Clements (2012).

[33] See also correlation analysis that is reported in Appendix B.4.

This may be due to a degree of multicollinearity between the IQR and the standard deviation of point forecasts and/or mean confidence intervals. In specification (c) the effect of the IQR is insignificant, while in specifications (a) and (b) we observe that only average individual confidence interval has a positive and highly significant effect on sd. Therefore, we can conclude that for policy maker to forecast inflation uncertainty in our framework it is most important to know the average individual confidence interval, which is still rarely the case in surveys of inflation opinions. These regressions reach similar conclusions to those from the survey data literature, e.g., Zarnowitz and Lambros (1987), Boero, Smith, and Wallis (2008), and Giordani and Söderlind (2003), who argue that average individual uncertainty is the proxy of inflation uncertainty that central banks should monitor.[34]

Inflation affects the standard deviation of inflation negatively, which might also be surprising. However, it is likely that if we separated the positive and negative developments of inflation we would find similar effects as in the above regressions for IQR and sdv, i.e. both terms would have significantly positive effects with negative development having a more profound effect. The output gap exerts a negative effect on sdv.

6 Conclusion

In this paper we have designed a macroeconomic experiment where subjects are asked to forecast inflation and its uncertainty. The underlying model of the economy is a simple NK model, which is commonly used for the analysis of monetary policy. The focus of the analysis has been on the confidence bounds reported by subjects as a measure of perceived uncertainty in the economy and in particular on the relationship with the monetary policy design. It has been shown that uncertainty has implications for both inflation outcomes and for unemployment and is an increasingly important indicator for monetary policy making. Similarly to inflation expectations, the formation of confidence bounds has also

[34] According to Boero, Smith, and Wallis (2012) the usefulness of a disagreement among point forecasts as an indicator seems to depend on the underlying macroeconomic environment. Also Lahiri and Sheng (2010) show that the aggregate forecast uncertainty can be decomposed to the disagreement among forecasters and the perceived variability of future aggregate shocks.

been found to be heterogeneous.

In different treatments we have focused on various modifications of the original Taylor rule and studied the influence of different monetary policy designs on the formation of confidence bounds. We have found that inflation targeting produces lower uncertainty and higher accuracy of intervals than inflation forecast targeting. The treatments with monetary policy that reacts strongly to deviations in inflation expectations from the inflation target also produce lower uncertainty and higher accuracy of confidence bounds, compared to treatments with monetary policy that does not react as strongly to deviations in inflation forecasts. This effect not only channels through the variability of inflation, but there is also evidence that there are additional effects.

Subjects on average underestimate risk. This is a common result in the psychology literature and is known as overconfidence bias. We have found find that only in 60.5% of cases do subjects correctly estimate risk. In particular, fewer than 10% of subjects on average report confidence bounds that approximately represent the 95% confidence intervals consistent with the actual realizations; around 10% overestimate risk, while all others underestimate risk.

We have also analyzed measures of individual uncertainty, disagreement among forecasters and the properties of aggregate distribution. We have found that confidence intervals are positively related to inflation variability, that they are highly inertial and that they widen after an "error." It is also interesting to observe the relation between inflation and confidence intervals. We have shown that below target inflation causes the interval to increase and absolute deviations from the inflation target is an appropriate variable to take into account. Furthermore, we have been able to establish some facts about the differences between the formation of lower and upper bounds. In particular, we have found that the upper bound is more sensitive to the stage of the business cycle while the lower bound exhibits significantly more inertia.

More generally, we have also studied the determinants of the choice of asymmetric interval. In our treatments Asym, subjects have the possibility of choosing an asymmetric confidence interval, while in treatments Sym they are restricted to symmetric intervals.

We have found that in only about 12.5% of cases subjects choose symmetric intervals when they have the possibility of choosing an asymmetric interval. Moreover, in treatments Asym more than 35% of subjects report higher upper bounds than the lower ones, while only about 5% of subjects show the opposite pattern. Symmetric intervals are more likely to be observed when the interest rate is high and less likely when inflation is below the target. Symmetric intervals are also more common when the output gap is positive and rising compared to the opposite stage of the business cycle.

What determines the evolution of the standard deviation of point forecasts and the interquartile range of the aggregate forecast distribution? We have documented that the interquartile range is more inertial than the standard deviation of point forecasts, while they both increase when inflation is below the target level. We have also compared the forecasting performance of these measures and observed that the interquartile range of the aggregate distribution is the one that has the highest correlation with the actual uncertainty. Nevertheless, regression analysis suggests that the average individual confidence interval is the only measure that consistently affects our forecasting specifications significantly. More central banks should design their surveys in such a way that each individual provides their whole distribution of forecasts or at least some measure of the uncertainty of their forecasts. In this sense it might be enough if they were asked for their symmetric confidence intervals as in our treatments Sym. Generally, this would greatly enhance the informativeness of these surveys as central banks would also receive a proxy for inflation forecast uncertainty.

References

Adam, K. (2007): "Experimental Evidence on the Persistence of Output and Inflation," *Economic Journal*, 117(520), 603–636.

Ang, A., G. Bekaert, and M. Wei (2007): "Do macro variables, asset markets, or surveys forecast inflation better?," *Journal of Monetary Economics*, 54(4), 1163–1212.

Anufriev, M., and C. Hommes (2012): "Evolution of Market Heuristics," *Knowledge Engineering Review*, 27(2), 255–271.

Arifovic, J., and T. J. Sargent (2003): "Laboratory Experiments with an Expectational Phillips Curve," in *Evolution and Procedures in Central Banking*, ed. by D. E. Altig, and B. D. Smith. Cambridge University Press, Cambridge.

Assenza, T., P. Heemeijer, C. Hommes, and D. Massaro (2011): "Individual Expectations and Aggregate Macro Behavior," CeNDEF Working Papers 11-01, Universiteit van Amsterdam, Center for Nonlinear Dynamics in Economics and Finance.

Bank, E. C. (2012): "Assessing the anchoring of longer-term inflation expectations," *ECB Monthly Bulletin*, July, 65–78.

Batchelor, R., and P. Dua (1996): "Empirical Measures of Inflation Uncertainty: A Cautionary Note," *Applied Economics*, 28(3), 333–41.

Ben-David, I., J. R. Graham, and C. R. Harvey (2010): "Managerial Miscalibration," NBER Working Papers 16215, National Bureau of Economic Research, Inc.

Bloom, N. (2009): "The Impact of Uncertainty Shocks," *Econometrica*, 77(3), 623–685.

Blundell, R., and S. Bond (1998): "Initial conditions and moment restrictions in dynamic panel data models," *Journal of Econometrics*, 87(1), 115–143.

Boero, G., J. Smith, and K. F. Wallis (2008): "Uncertainty and disagreement in economic prediction: the Bank of England Survey of External Forecasters," *Economic Journal*, 118(530), 1107–27.

——— (2012): "The measurment and characteristics of professional forecasters' uncertainty," Mimeo, University of Warwick.

Bomberger, W. A. (1996): "Disagreement as a Measure of Uncertainty," *Journal of Money, Credit and Banking*, 28(3), 381–92.

Bottazzi, G., and G. Devetag (2005): "Expectations structure in asset pricing experiments," in *Nonlinear Dynamics and Heterogeneous Interacting Agents*, ed. by T. Lux, S. Reitz, and E. Samanidou, Lecture Notes in Economics and Mathematical Systems. Springer Verlag, Berlin, 550 edn.

Bruine de Bruin, W., C. F. Manski, G. Topa, and W. van der Klaauw (2011): "Measuring consumer uncertainty about future inflation," *Journal of Applied Econometrics*, 26(3), 454–478.

Capistrán, C., and M. Ramos-Francia (2010): "Does Inflation Targeting Affect the Dispersion of Inflation Expectations?," *Journal of Money, Credit and Banking*, 42(1), 113–134.

Capistrán, C., and A. Timmermann (2009): "Disagreement and Biases in Inflation Expectations," *Journal of Money, Credit and Banking*, 41(2-3), 365–396.

Carroll, C. D. (2003): "Macroeconomic Expectations Of Households and Professional Forecasters," *The Quarterly Journal of Economics*, 118(1), 269–298.

Clements, M. P. (2012): "Characterising how surveys of US macrovariables add value," Mimeo, University of Warwick, Department of Economics.

Cooley, T. F., and E. C. Prescott (1995): "Economic Growth and Business Cycles," in *Frontiers of Business Cycle Research*, ed. by T. F. Cooley. Princeton University Press, Princeton.

Cukierman, A., and P. Wachtel (1979): "Differential Inflationary Expectations and the Variability of the Rate of Inflation: Theory and Evidence," *American Economic Review*, 69(4), 595–609.

D'Amico, S., and A. Orphanides (2008): "Uncertainty and disagreement in economic forecasting," Finance and Economics Discussion Series 2008-56, Board of Governors of the Federal Reserve System (U.S.).

Diebold, F. X., J. Hahn, and A. S. Tay (1999): "Multivariate Density Forecast Evaluation And Calibration In Financial Risk Management: High-Frequency Returns On Foreign Exchange," *The Review of Economics and Statistics*, 81(4), 661–673.

Du, N., and D. V. Budescu (2007): "Does past volatility affect investors' price forecasts and confidence judgements?," *International Journal of Forecasting*, 23(3), 497–511.

Duffy, J. (2008): "Macroeconomics: A Survey of Laboratory Research," Working Papers 334, University of Pittsburgh, Department of Economics.

Engelberg, J., C. F. Manski, and J. Williams (2009): "Comparing the Point Predictions and Subjective Probability Distributions of Professional Forecasters," *Journal of Business & Economic Statistics*, 27, 30–41.

Evans, G. W., and S. Honkapohja (2001): *Learning and Expectations in Macroeconomics*. Princeton University Press.

Evans, G. W., S. Honkapohja, and R. Marimon (2001): "Convergence in monetary inflation models with heterogeneous learning rules," *Macroeconomic Dynamics*, 5(1), 1–31.

Evans, M., and P. Wachtel (1993): "Inflation Regimes and the Sources of Inflation Uncertainty," *Journal of Money, Credit and Banking*, 25(3), 475–511.

Fehr, E., and J.-R. Tyran (2008): "Limited Rationality and Strategic Interaction: The Impact of the Strategic Environment on Nominal Inertia," *Econometrica*, 76(2), 353–394.

Fischbacher, U. (2007): "z-Tree: Zurich toolbox for ready-made economic experiments," *Experimental Economics*, 10(2), 171–178.

Friedman, M. (1968): "The role of monetary policy," *American Economic Review*, 58(1), 1–17.

——— (1977): "Nobel Lecture: Inflation and Unemployment," *Journal of Political Economy*, 85(3), 451–72.

García, J. A., and A. Manzanares (2007): "What can probability forecasts tell us about inflation risks?," Working Paper Series 825, European Central Bank.

Gilovich, T., D. Griffin, and D. Kahneman (2002): *Heuristics and Biases: The Psychology of Intuitive Judgment*. Cambridge University Press, Cambridge.

Giordani, P., and P. Söderlind (2003): "Inflation forecast uncertainty," *European Economic Review*, 47(6), 1037–1059.

——— (2006): "Is there evidence of pessimism and doubt in subjective distributions? Implications for the equity premium puzzle," *Journal of Economic Dynamics and Control*, 30(6), 1027–1043.

Hoffrage, U. (2004): "Overconfidence," in *Cognitive illusions: A handbook on fallacies and biases in thinking, judgement and memory*, ed. by R. F. Pohl, pp. 235–254. Psychology Press.

Hommes, C. (2011): "The heterogeneous expectations hypothesis: Some evidence from the lab," *Journal of Economic Dynamics and Control*, 35(1), 1–24.

Hommes, C., J. Sonnemans, J. Tuinstra, and H. van de Velden (2005): "Coordination of Expectations in Asset Pricing Experiments," *Review of Financial Studies*, 18(3), 955–980.

Ireland, P. N. (2004): "Technology Shocks in the New Keynesian Model," *The Review of Economics and Statistics*, 86(4), 923–936.

Keren, G. (1991): "Calibration and probability judgements: Conceptual and methodological issues," *Acta Psychologica*, 77(3), 217 – 273.

Lahiri, K., and X. Sheng (2010): "Measuring forecast uncertainty by disagreement: The missing link," *Journal of Applied Econometrics*, 25(4), 514–538.

Lawrence, M., and S. Makridakis (1989): "Factors affecting judgmental forecasts and confidence intervals," *Organizational Behavior and Human Decision Processes*, 43(2), 172–187.

Lawrence, M., and M. O'Connor (1992): "Exploring judgemental forecasting," *International Journal of Forecasting*, 8(1), 15–26.

Levi, M. D., and J. H. Makin (1980): "Inflation Uncertainty and the Phillips Curve: Some Empirical Evidence," *American Economic Review*, 70(5), 1022–27.

Lichtenstein, S., B. Fischhoff, and L. D. Phillips (1982): "Calibration of probabilities: the state of the art to 1980," in *Judgment under uncertainty: heuristics and biases*, ed. by D. Kahneman, P. Slovic, and A. Tversky, NBER Chapters, pp. 306–334. Cambridge University Press, Cambridge, UK.

Liu, F., and K. Lahiri (2006): "Modelling multi-period inflation uncertainty using a panel of density forecasts," *Journal of Applied Econometrics*, 21(8), 1199–1219.

Lorenzoni, G. (2010): "Optimal Monetary Policy with Uncertain Fundamentals and Dispersed Information," *Review of Economic Studies*, 77(1), 305–338.

Mankiw, N. G., R. Reis, and J. Wolfers (2004): "Disagreement about Inflation Expectations," *NBER Macroeconomics Annual 2003*, 18, 209–248.

Marcet, A., and T. J. Sargent (1989): "Convergence of Least-Squares Learning in Environments with Hidden State Variables and Private Information," *Journal of Political Economy*, 97(6), 1306–22.

Marimon, R., S. E. Spear, and S. Sunder (1993): "Expectationally Driven Market Volatility: An Experimental Study," *Journal of Economic Theory*, 61(1), 74–103.

McCallum, B. T., and E. Nelson (1999): "Performance of Operational Policy Rules in an Estimated Semi-Classical Model," in *Monetary Policy Rules*, ed. by J. Taylor, pp. 15–45. University of Chicago Press.

Mishkin, F. S. (2008): "Whither Federal Reserve Communications," *Speech at the Peterson Institute for International Economics*, Washington, D.C., July 28, 2008.

Mullineaux, D. J. (1980): "Unemployment, Industrial Production, and Inflation Uncertainty in the United States," *The Review of Economics and Statistics*, 62(2), 163–69.

Muradoglu, G., and D. Onkal (1994): "An exploratory analysis of portfolio managers' probabilistic forecasts of stock prices," *Journal of Forecasting*, 13(7), 565–578.

Noussair, C. N., D. Pfajfar, and J. Zsiros (2011): "Frictions, persistence, and central bank policy in an experimental dynamic stochastic general equilibrium economy," Discussion Paper 2011-030, Tilburg University, Center for Economic Research.

Oskamp, S. (1965): "Overconfidence in case-study judgements," *The Journal of Consulting Psychology*, 29(3), 261–265.

Patton, A. J., and A. Timmermann (2008): "The Resolution of Macroeconomic Uncertainty: Evidence from Survey Forecast," CREATES Research Papers 2008-54, School of Economics and Management, University of Aarhus.

Pfajfar, D., and E. Santoro (2010): "Heterogeneity, learning and information stickiness in inflation expectations," *Journal of Economic Behavior & Organization*, 75(3), 426–444.

Pfajfar, D., and B. Žakelj (2011): "Inflation Expectations and Monetary Policy Design: Evidence from the Laboratory," Discussion Paper 2011-091, Tilburg University, Center for Economic Research.

Rich, R., and J. Tracy (2010): "The Relationships among Expected Inflation, Disagreement, and Uncertainty: Evidence from Matched Point and Density Forecasts," *The Review of Economics and Statistics*, 92(1), 200–207.

Rich, R. W., and J. S. Butler (1998): "Disagreement as a Measure of Uncertainty: A Comment on Bomberg," *Journal of Money, Credit and Banking*, 30(3), pp. 411–419.

Söderlind, P. (2011): "Inflation Risk Premia and Survey Evidence on Macroeconomic Uncertainty," *International Journal of Central Banking*, 7(2), 113–133.

Thaler, R. H. (2000): "From Homo Economicus to Homo Sapiens," *Journal of Economic Perspectives*, 14(1), 133–141.

Van Nieuwerburgh, S., and L. Veldkamp (2006): "Learning asymmetries in real business cycles," *Journal of Monetary Economics*, 53(4), 753–772.

Wallis, K. F. (2004): "An assessment of Bank of England and National Institute inflation forecast uncertainties," *National Institute Economic Review*, 189, 64–71.

——— (2005): "Combining density and interval forecasts: a modest proposal," *Oxford Bulletin of Economics and Statistics*, 67, 983–994.

Walsh, C. E. (2003): *Monetary Theory and Policy*. The MIT Press, second edn.

Williams, A. W. (1987): "The Formation of Price Forecasts in Experimental Markets," *Journal of Money, Credit and Banking*, 19(1), 1–18.

Woodford, M. (2003): *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton University Press.

Zarnowitz, V., and L. A. Lambros (1987): "Consensus and Uncertainty in Economic Prediction," *Journal of Political Economy*, 95(3), 591–621.

Figures and Tables

policy regime p		Treatments Sym-p Symmetric confidence interval	Treatments Asym-p Asymmetric confidence bounds
Taylor rule (equation)	Parameters	Groups	Groups
1 – Forward looking (3)	$\gamma = 1.5$	1-4	5-6
2 – Forward looking (3)	$\gamma = 1.35$	7-10	11-12
3 – Forward looking (3)	$\gamma = 4$	13-16	17-18
4 – Contemporaneous (4)	$\gamma = 1.5$	19-22	23-24

Table 1: Treatments

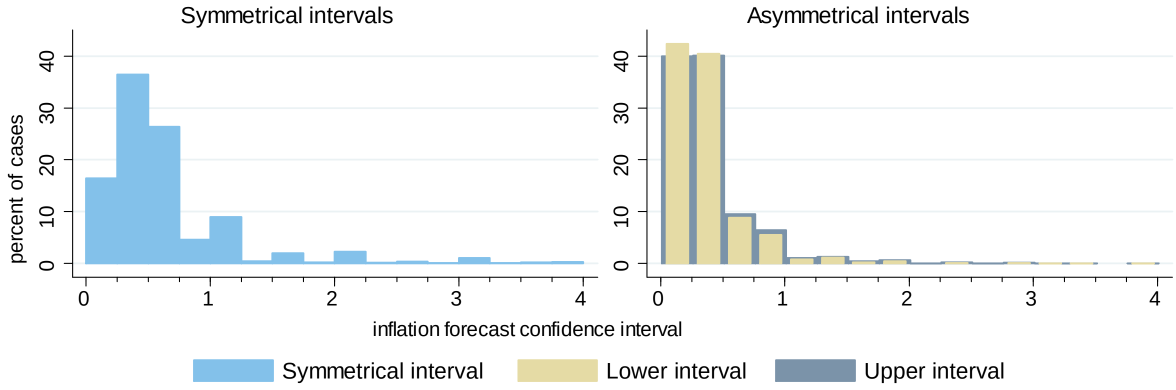

Figure 1: Histogram of confidence intervals for all treatments, subjects and periods

Average confidence interval	All	Treat. Sym (symmetric)	Treat. Asym (asymmetric)
1 – Forward looking (3), $\gamma = 1.5$	0.564	0.669	0.352
2 – Forward looking (3), $\gamma = 1.35$	0.776	0.914	0.500
3 – Forward looking (3), $\gamma = 4$	0.395	0.466	0.254
4 – Contemporaneous (4), $\gamma = 1.5$	0.430	0.410	0.471

Table 2: Width of confidence intervals across treatments. Note: The width of asymmetric confidence intervals is calculated as (Upper b. - Lower b.)/2.

		Inflation		Confidence bound					
				Symmetric		Lower		Upper	
Treat.	Group	mean	stdev	mean	stdev	mean	stdev	mean	stdev
Sym-1	1	2.85	5.87	0.97	0.71	-	-	-	-
Sym-1	2	2.88	2.91	0.65	0.40	-	-	-	-
Sym-1	3	2.92	1.97	0.70	0.35	-	-	-	-
Sym-1	4	3.00	0.76	0.34	0.16	-	-	-	-
Aysm-1	5	3.13	1.10	-	-	0.36	0.19	0.41	0.24
Aysm-1	6	3.12	0.90	-	-	0.29	0.14	0.35	0.41
Sym-2	7	3.12	0.76	1.09	0.30	-	-	-	-
Sym-2	8	3.09	1.82	1.15	0.63	-	-	-	-
Sym-2	9	3.13	0.51	0.38	0.21	-	-	-	-
Sym-2	10	3.02	5.53	1.02	0.56	-	-	-	-
Aysm-2	11	2.52	3.58	-	-	0.61	0.45	0.72	0.43
Aysm-2	12	3.03	0.88	-	-	0.33	0.12	0.33	0.14
Sym-3	13	3.01	0.52	0.53	0.13	-	-	-	-
Sym-3	14	3.02	0.94	0.65	0.32	-	-	-	-
Sym-3	15	2.99	0.24	0.35	0.09	-	-	-	-
Sym-3	16	3.00	0.26	0.33	0.10	-	-	-	-
Aysm-3	17	2.99	0.31	-	-	0.28	0.09	0.28	0.10
Aysm-3	18	3.01	0.24	-	-	0.20	0.08	0.25	0.35
Sym-4	19	3.09	0.39	0.36	0.13	-	-	-	-
Sym-4	20	3.23	0.81	0.56	0.20	-	-	-	-
Sym-4	21	3.05	0.48	0.38	0.09	-	-	-	-
Sym-4	22	3.05	0.38	0.34	0.10	-	-	-	-
Aysm-4	23	3.09	0.52	-	-	0.31	0.12	0.31	0.15
Aysm-4	24	3.11	1.29	-	-	0.60	0.28	0.65	0.37
All Sym		3.03	1.51	0.61	0.28	-	-	-	-
All Asym		3.00	1.10	-	-	0.37	0.18	0.41	0.28

Table 3: Confidence bounds, summary statistics.

$sip^k_{t+1\|t}$:	all	$treat.Sym$	$treat.Asym-L$	$treat.Asym-U$
$sip^k_{t\|t-1}$	0.4390***	0.5445***	0.4407***	0.0925
	(0.1114)	(0.0921)	(0.0485)	(0.0982)
sd^j_{t-1}	0.1167***	0.0955**	0.1357***	0.2643***
	(0.0450)	(0.0401)	(0.0220)	(0.0561)
α	0.2143***	0.2039***	0.1142***	0.1884***
	(0.0283)	(0.0285)	(0.0187)	(0.0323)
N	14904	9936	4968	4968
Wald $\chi^2_{(3)}$	140.9	259.1	346.1	34.6

Table 4: Confidence intervals and standard deviation of inflation. Note: coefficients are based on the Blundell-Bond system GMM estimator. Standard errors in parentheses are calculated using bootstrap procedures (1000 replications) that take into account the potential presence of clusters in groups. */**/*** denotes significance at 10/5/1 percent level.

$sip^k_{t+1\|t}$:	*all*	*treat.Sym*	*treat.Asym − L*	*treat.Asym − U*
$sip^k_{t\|t-1}$	0.3976***	0.5333***	0.4305***	0.0900
	(0.1034)	(0.0990)	(0.0398)	(0.0997)
D_1y_{t-1}	0.0067	0.0198	0.0202	-0.0560
	(0.0219)	(0.0258)	(0.0252)	(0.0465)
D_2y_{t-1}	-0.0188	-0.0118	-0.0144	-0.0650**
	(0.0225)	(0.0217)	(0.0304)	(0.0262)
D_3y_{t-1}	0.0067	0.0183	0.0051	-0.1142***
	(0.0296)	(0.0271)	(0.0183)	(0.0413)
i_{t-1}	0.0110	0.0070	-0.0066	0.0025
	(0.0076)	(0.0073)	(0.0059)	(0.0157)
$D_L\|\pi_{t-1}\|$	0.0294**	0.0241**	0.0247**	0.0782***
	(0.0115)	(0.0105)	(0.0108)	(0.0234)
$D_H\|\pi_{t-1}\|$	0.0180	0.0173	0.0668***	0.0248
	(0.0167)	(0.0135)	(0.0139)	(0.0310)
$\|r^k_{t-1}\|$	0.0552***	0.0473***	0.0474**	0.0749***
	(0.0154)	(0.0159)	(0.0203)	(0.0250)
$T2$	1.0505*			
	(0.5519)			
$T3$	-0.6098			
	(0.5743)			
$T4$	-0.6351			
	(0.5913)			
α	0.2790	0.2062***	0.1694***	0.2898***
	(0.2893)	(0.0415)	(0.0266)	(0.0541)
N	14688	9792	4896	4896
Wald $\chi^2_{(12)}$	393.8	715.8	865.0	145.0

Table 5: Confidence intervals and macroeconomic variables. Note: $treat.B-L$ ($treat.B-U$) only includes the part of the interval beneath (above) the point forecast. Coefficients are based on the Blundell-Bond system GMM estimator. Standard errors in parentheses are calculated using bootstrap procedures (1000 replications) that take into account the potential presence of clusters in groups. */**/*** denotes significance at 10/5/1 percent level.

Lower vs. upper (% of subjects)	$C_L < C_U$	$C_L \approx C_U$	$C_L > C_U$
1 – Forward looking (3), $\gamma = 1.5$	44.4	50.0	5.6
2 – Forward looking (3), $\gamma = 1.35$	50.0	44.4	5.6
3 – Forward looking (3), $\gamma = 4$	33.3	66.7	0.0
4 – Contemporaneous (4), $\gamma = 1.5$	16.7	72.2	11.1
All	36.1	58.3	5.6

Table 6: Proportions of subjects from treatments Asym, depending on the difference between their upper (C_U) and lower (C_L) confidence intervals. When $C_L < C_U$, the subject chose on average a smaller lower interval than upper interval. Based on pairwise t-test with 5% signifficance level.

	Symmetry		Skewness
	D_7	D_8	skw_t^k
skw_{t-1}^k	-	-	0.2861***
			(0.0576)
$sip_{t\|t-1}^k$	0.2498	-0.7167	-0.2375***
	(0.1643)	(0.5136)	(0.0852)
$D_1 y_{t-1}$	0.3345	0.4867**	-0.0496
	(0.4229)	(0.2048)	(0.0415)
$D_2 y_{t-1}$	-0.4418**	0.1259	-0.0447
	(0.2093)	(0.2308)	(0.0337)
$D_3 y_{t-1}$	-0.3388	0.1152***	-0.0776***
	(0.2504)	(0.0420)	(0.0240)
i_{t-1}	0.2547*	0.2111***	0.0004
	(0.1306)	(0.0684)	(0.0150)
$D_L\|\pi_{t-1}\|$	0.1828	0.1802	0.0273
	(0.2757)	(0.1174)	(0.0275)
$D_H\|\pi_{t-1}\|$	-0.4488*	-0.2613**	-0.0126
	(0.2550)	(0.1177)	(0.0232)
sd_{t-1}^k	-0.3237**	-0.5066	-0.0050
	(0.1510)	(0.3272)	(0.0498)
α	-	-	0.1037**
			(0.0519)
N	4968	4968	4968
Wald $\chi^2_{(8,9)}$	79.3	58.3	156.3

Table 7: Determinants of symmetric and skewed intervals. Note: coefficients for the symmetry tests are based on fixed effects logit estimations, while coefficients for skewness are based on the Blundell-Bond system GMM estimator. Standard errors in parentheses are calculated using bootstrap procedures (1000 replications) that take into account the potential presence of clusters in groups. */**/*** denotes significance at 10/5/1 percent level.

	$sdv^j_{t+1\|t}$			$IQR^j_{t+1\|t}$		
	all	*treat.Sym*	*treat.Asym*	*all*	*treat.Sym*	*treat.Asym*
$sdv^j_{t\|t-1}$	0.1463	0.1265	0.4970***			
	(0.1409)	(0.1046)	(0.0247)			
$IQR^j_{t\|t-1}$				0.4982***	0.4738***	0.6280***
				(0.0896)	(0.0787)	(0.0670)
$D_1 y_{t-1}$	-0.0171	0.0122	0.0157	-0.0298	-0.0282	0.0385
	(0.0154)	(0.0168)	(0.0376)	(0.0622)	(0.0854)	(0.0582)
$D_2 y_{t-1}$	0.0026	0.0136	-0.1275***	-0.0809	-0.0713	-0.1122**
	(0.0311)	(0.0250)	(0.0164)	(0.0492)	(0.0572)	(0.0475)
$D_3 y_{t-1}$	0.0392	0.0520	-0.0249	0.0848	0.1073	-0.0538
	(0.0593)	(0.0749)	(0.0369)	(0.0841)	(0.1000)	(0.0402)
i_{t-1}	0.0279	0.0249	-0.0002	0.0083	0.0076	0.0109
	(0.0330)	(0.0288)	(0.0389)	(0.0131)	(0.0210)	(0.0246)
$D_L\|\pi_{t-1}\|$	0.1430***	0.1533***	0.0773	0.0758**	0.0789***	0.0497
	(0.0507)	(0.0353)	(0.0534)	(0.0294)	(0.0286)	(0.0345)
$D_H\|\pi_{t-1}\|$	0.0787	0.0901	0.0794	0.0438	0.0492	0.0022
	(0.0717)	(0.0701)	(0.0675)	(0.0530)	(0.0615)	(0.0401)
mr^j_{t-1}	0.2211***	0.2447***	0.0704	0.2174***	0.2438***	0.0790
	(0.0297)	(0.0169)	(0.0779)	(0.0348)	(0.0184)	(0.0959)
α	-0.0218	-0.0252	0.0332	0.0739**	0.0836	0.0668
	(0.0911)	(0.0726)	(0.1211)	(0.0308)	(0.0610)	(0.0750)
N	1632	1088	544	1632	1088	544
Wald $\chi^2_{(8)}$	3763.3	12747.1	5495.4	19215.1	15228.4	3032.3

Table 8: Analysis of Disagreement: Interquartile Range (left) and Standard Deviation of Point Forecasts (right). Note: coefficients are based on the Blundell-Bond system GMM estimator. Standard errors in parentheses are calculated using bootstrap procedures (1000 replications) that take into account the potential presence of clusters in treatments. */**/*** denotes significance at 10/5/1 percent level.

sd_t^j :	(a)	(b)	(c)
sd_{t-1}^j	0.9913***	0.9843***	1.0036***
	(0.0124)	(0.0116)	(0.0137)
$asip_{t-1}^j$	0.0837***	0.0708**	-
	(0.0298)	(0.0289)	
sdv_{t-1}^j	0.0106	0.0073	-
	(0.0179)	(0.0152)	
IQR_{t-1}^j	-0.0170*	-	-0.0018
	(0.0100)		(0.0114)
i_{t-1}	0.0108	0.0108	0.0129
	(0.0084)	(0.0082)	(0.0096)
π_{t-1}	-0.0135	-0.0136	-0.0148
	(0.0118)	(0.0115)	(0.0137)
y_{t-1}	-0.0094*	-0.0092*	-0.0125***
	(0.0052)	(0.0049)	(0.0037)
α	-0.0109	-0.0071	0.0169
	(0.0227)	(0.0218)	(0.0178)
N	1656	1656	1656
Wald $\chi^2_{(7,6,5)}$	54840.3	50525.4	22529.2

Table 9: Factors affecting the standard deviation of inflation. Note: coefficients are based on the Blundell-Bond system GMM estimator. Standard errors in parentheses are calculated using bootstrap procedures (1000 replications) that take into account the potential presence of clusters in treatments. */**/*** denotes significance at 10/5/1 percent level.

A Additional Tables and Figures

		Confidence bound								
		Symmetric			Lower			Upper		
Treat	Group	$<$	$\approx$	$>$	$<$	$\approx$	$>$	$<$	$\approx$	$>$
Sym-1	1	100	0	0	-	-	-	-	-	-
Sym-1	2	78	11	11	-	-	-	-	-	-
Sym-1	3	89	11	0	-	-	-	-	-	-
Sym-1	4	78	22	0	-	-	-	-	-	-
Aysm-1	5	-	-	-	89	11	0	0	22	78
Aysm-1	6	-	-	-	100	0	0	0	11	89
Sym-2	7	44	11	44	-	-	-	-	-	-
Sym-2	8	78	11	11	-	-	-	-	-	-
Sym-2	9	100	0	0	-	-	-	-	-	-
Sym-2	10	100	0	0	-	-	-	-	-	-
Aysm-2	11	-	-	-	100	0	0	0	0	100
Aysm-2	12	-	-	-	100	0	0	0	0	100
Sym-3	13	56	22	22	-	-	-	-	-	-
Sym-3	14	89	11	0	-	-	-	-	-	-
Sym-3	15	56	11	33	-	-	-	-	-	-
Sym-3	16	100	0	0	-	-	-	-	-	-
Aysm-3	17	-	-	-	100	0	0	0	0	100
Aysm-3	18	-	-	-	100	0	0	0	11	89
Sym-4	19	78	11	11	-	-	-	-	-	-
Sym-4	20	89	11	0	-	-	-	-	-	-
Sym-4	21	67	0	33	-	-	-	-	-	-
Sym-4	22	78	11	11	-	-	-	-	-	-
Aysm-4	23	-	-	-	78	0	22	11	11	78
Aysm-4	24	-	-	-	100	0	0	0	11	89
All		80	9	11	96	1	3	1	8	90

Table A1: Percentage of subjects by group with underprediction/overprediction of confidence interval. Note: the benchmark confidence level is $1.96 * sd^k_{t-1}$. $<$ ($>$) identifies frequencies of subjects whose inputs are signifficantly lower (higher) than the benchmark value. $\approx$ identifies subjects whose input is not siggnificantly different from the benchmark. Based on t-tests.

	All			treatments Sym			treatments Asym		
Inflation	$\uparrow$	$\downarrow$	$\sim$	$\uparrow$	$\downarrow$	$\sim$	$\uparrow$	$\downarrow$	$\sim$
Underprediction	34.63	3.98	17.83	30.93	4.12	16.79	41.39	3.69	20.02
Inside interval	60.65	58.41	63.95	64.81	62.75	66.43	53.03	49.15	58.76
Overprediction	4.72	37.6	18.22	4.25	33.13	16.79	5.58	47.17	21.22

Table A2: Interval correctness depending on the phase of the inflation cycle (% of decisions). $\uparrow$ denotes cases when inflation increases for at least the last 2 periods, and $\downarrow$ denotes cases when it decreases for at least the last 2 periods. $\sim$ represents all other cases. Subjects "underpredict" when the actual inflation is larger than their predicted upper confidence bound; and "overpredict" when the actual inflation is lower than their predicted lower confidence bound.

Monetary Policy Regime	Groups	Mean standard deviation	Median standard deviation	Comparison with treat. 1 (p-value)
1: Inf. forc. targ. $\gamma = 1.5$	$1-6$	2.24	1.52	–
2: Inf. forc. targ. $\gamma = 1.35$	$7-12$	2.17	1.35	0.6310
3: Inf. forc. targ. $\gamma = 4$	$13-18$	0.42	0.29	0.0104
4: Inf. targeting $\gamma = 1.5$	$19-24$	0.64	0.50	0.0250

Table A3: Standard deviation of inflation for each treatment and Kruskal-Wallis test of differnences between treatments using group-level standard deviations.

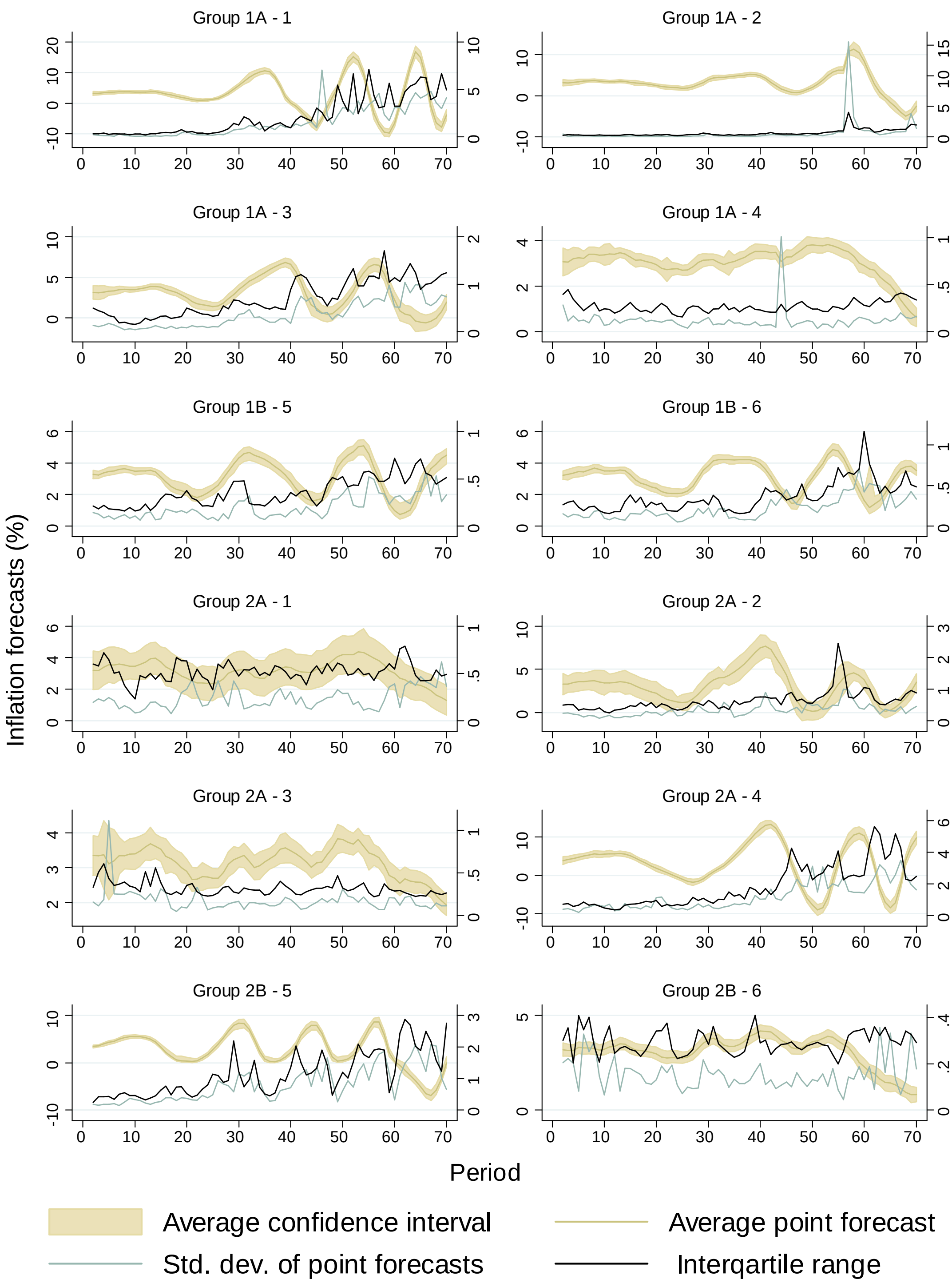

Figure A1: Average inflation forecasts and average confidence intervals (left axis) and disagreement and uncertainty measures (right axis) per group. Interquartile range is calculated from the aggregate expectation distribution as described in Section 4.1.

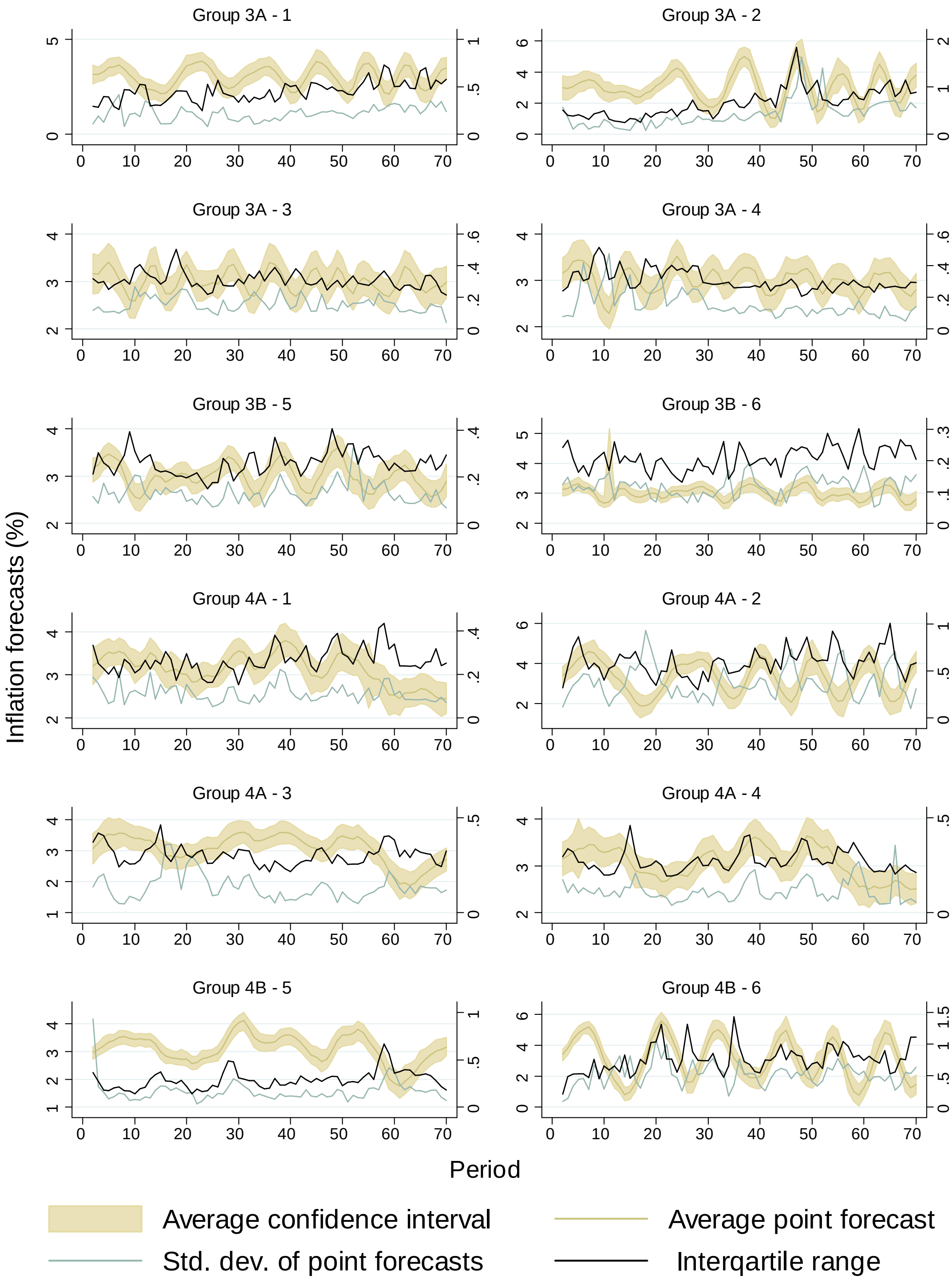

Figure A2: Average inflation forecasts and average confidence intervals (left axis) and disagreement and uncertainty measures (right axis) per group. Interquartile range is calculated from the aggregate expectation distribution as described in Section 4.1.

B Additional analysis (not for publication)

B.1 Forecasting accuracy

We check how the volatility of inflation, the width of confidence bounds, and macroeconomic variables affect the likelihood of inflation falling within the specified confidence bound. We estimate the following regression, where x_t^k takes value 1 when inflation falls within the provided bounds and 0 otherwise:

$$x_t^k = \alpha^k + \beta sip_{t|t-1}^k + \gamma D_1 y_{t-1} + \delta D_2 y_{t-1} + \epsilon D_3 y_{t-1} + \tag{12}$$
$$+\eta D_L \left|\pi_{t-1}\right| + \theta D_H \left|\pi_{t-1}\right| + \zeta i_{t-1} + \delta sd_{t-1}^j + u_t^{em},$$

x_t^k :	*all*	*treat.Sym*	*treat.Asym*
$sip_{t\|t-1}^k$	2.3985***	2.3578***	2.9678***
	(0.2340)	(0.3620)	(0.4567)
$D_1 y_{t-1}$	-0.8720***	-1.1590***	-0.7103***
	(0.2117)	(0.4328)	(0.2137)
$D_2 y_{t-1}$	1.3565***	1.9346***	1.4602***
	(0.2304)	(0.5309)	(0.2439)
$D_3 y_{t-1}$	0.3092*	0.3000	0.2717
	(0.1684)	(0.3153)	(0.2023)
$D_L\|\pi_{t-1}\|$	0.2179**	0.0933	0.3218*
	(0.0948)	(0.5938)	(0.1856)
$D_H\|\pi_{t-1}\|$	0.5955***	1.2236**	0.5659***
	(0.1344)	(0.4821)	(0.1497)
i_{t-1}	-0.1529**	-0.3655	-0.0960
	(0.0758)	(0.3859)	(0.0817)
sd_{t-1}^j	-1.4642***	-0.8690*	-1.8730***
	(0.2722)	(0.4525)	(0.4803)
N	14628	4968	9660
Wald $\chi^2_{(8)}$	168.4	230.0	122.9

Table B1: Forecasting accuracy and confidence intervals. Note: coefficients are based on fixed effects logit estimations. Standard errors in parentheses are calculated using bootstrap procedures (1000 replications) that take into account the potential presence of clusters in groups. */**/*** denotes significance at 10/5/1 percent level.

The results for fixed effects logit estimation are reported in Table B1 while those for Poisson fixed effects and random effects are reported in the working paper version (results remain virtually unchanged). As one would expect, when there is a higher volatility of inflation there are more results outside the interval, especially in treatments Asym. This is well documented in the psychology literature as greater volatility leads to overconfidence (e.g., Lawrence and Makridakis, 1989, Lawrence and O'Connor, 1992).[35] However, some studies also find that there is no such effect (Du and Budescu, 2007). In both treatments wider confidence intervals result in a higher probability of correctly specifying the confidence interval. Interestingly, we can observe that there exists some pattern across business cycles. There are more outcomes outside the interval, when the output gap is positive and has a clear upward trend of inflation, while in the opposite situation there is a lower probability of misperceiving inflation uncertainty. Inflation also has a significant positive impact on the likelihood of the forecast falling within the interval, especially when inflation is above the target value.

In Table B2 we also report the results of the relationship between the individual k's forecast error $r^k_{t+1} = \pi^k_{t+1|t} - \pi_{t+1}$, and the confidence interval as a measure of uncertainty:

$$r^k_{t+1} = \alpha + \beta r^k_t + \gamma sip^k_{t+1|t} + u^{em}_t. \tag{13}$$

r^k_{t+1} :	all	$treat.Sym$	$treat.Asym - L$	$treat.Asym - U$
r^k_t	0.6970***	0.6757***	0.8521***	0.8524***
	(0.1376)	(0.1596)	(0.0250)	(0.0254)
$sip^k_{t+1\|t}$	0.0559	0.0812	0.6387***	-0.1139*
	(0.0928)	(0.1102)	(0.1980)	(0.0619)
α	-0.0211	-0.0401	-0.2016***	-0.0076
	(0.0513)	(0.0651)	(0.0468)	(0.0299)
N	14688	9792	4896	4896
Wald $\chi^2_{(3)}$	26.7	21.4	6625.1	4809.6

Table B2: Forecast errors and confidence intervals. Note: coefficients are based on the Blundell-Bond system GMM estimator. Standard errors in parentheses are calculated using bootstrap procedures (1000 replications) that take into account the potential presence of clusters in groups. */**/*** denotes significance at 10/5/1 percent level.

[35] Psychologists argue that this overconfidence is due to hard-easy effects, i.e. miscalibration (reported narrower confidence intervals) is higher in hard tasks and attenuated or even eliminated in easy tasks (e.g. Keren, 1991).

B.2 Determinants of individual confidence intervals

A second feature of the confidence intervals that we want to study is the subjects' responses to inflation falling outside the confidence interval. To discriminate between the effects of overshooting and undershooting we introduce two dummy variables. D_4^k takes value 1 if $\left(\left|r_{t-1}^k\right| > sip_{t-1}^k\right) \wedge \left(r_{t-1}^k \geq 0\right)$, and 0 otherwise. Note that $r_{t-1}^k = \pi_{t-1} - \pi_{t-1|t-2}^k$ is subject k's last observed forecast error. D_5^k equals 1 if $\left(\left|r_{t-1}^k\right| > sip_{t-1}^k\right) \wedge \left(r_{t-1}^k \leq 0\right)$, and 0 otherwise, while D_6^k is 1 when $\left|r_{t-1}^k\right| < sip_{t-1}^k$, and 0 otherwise. Therefore $D_4^k = 1$ when subject k underestimates inflation; while $D_5^k = 1$ when subject k overestimates inflation. We run the following regression:

$$sip_{t+1|t}^k = \alpha + \beta sip_{t|t-1}^k + \gamma D_4^k r_{t-1}^k + \delta D_5^k r_{t-1}^k + \epsilon D_6^k r_{t-1}^k + u_t^{em}. \tag{14}$$

$sip_{t+1\|t}^k$:	all	$treat.Sym$	$treat.Asym-L$	$treat.Asym-U$
$sip_{t\|t-1}^k$	0.4430***	0.5496***	0.4641***	0.1068
	(0.1080)	(0.0865)	(0.0491)	(0.1059)
$D_4 r_{t-1}^k$	0.0363**	0.0292**	0.0023	0.0669*
	(0.0153)	(0.0147)	(0.0228)	(0.0343)
$D_5 r_{t-1}^k$	-0.0760***	-0.0647***	-0.0955***	-0.0668**
	(0.0193)	(0.0190)	(0.0094)	(0.0269)
$D_6 r_{t-1}^k$	0.0015	0.0025	-0.0191*	0.0506
	(0.0201)	(0.0204)	(0.0107)	(0.0309)
α	0.2799***	0.2568***	0.1882***	0.3504***
	(0.0396)	(0.0416)	(0.0251)	(0.0406)
N	14688	9792	4896	4896
Wald $\chi^2_{(5)}$	203.6	248.5	1048.1	19.7

Table B3: Confidence intervals and phases of the economic cycle. Note: $treat.B-L$ ($treat.B-U$) only includes part of the interval beneath (above) the point forecast. Coefficients are based on the Blundell-Bond system GMM estimator. Standard errors in parentheses are calculated using bootstrap procedures (1000 replications) that take into account the potential presence of clusters in groups. */**/*** denotes significance at 10/5/1 percent level.

The results shown in Table B3 suggest that subjects increased their confidence intervals after the last observed inflation is outside the interval.[36] This holds for both "undershooting" and "overshooting." In the latter case r^k_{t-1} is negative, so a negative coefficient δ implies that confidence intervals are widened after $\left|r^k_{t-1}\right| > sip^k_{t-1}$. Positive or negative errors do not result in any significant change in confidence intervals in the next period when inflation falls within the interval. It is also interesting to note that the confidence intervals in treatments Asym exhibit less inertia, especially at the upper bound, compared to treatments Sym. Moreover, the interval above the point forecast widens with both overshooting and undershooting while the interval below is more stable and responds only to undershooting. This also represents the first potential source of observed asymmetries. Ben-David, Graham, and Harvey (2010) also note that there is a difference regarding the formation of the upper and the lower bound of confidence intervals. They argue that lower forecast bounds are significantly affected by the past return while upper ones are not.

Additionally, we also report robustness of the results in the text in the next 3 tables.

$sip^k_{t+1\|t}$:	all	$treat.Sym$	$treat.Asym - L$	$treat.Asym - U$
$sip^k_{t\|t-1}$	0.4472***	0.5530***	0.4468***	0.0991
	(0.1058)	(0.0817)	(0.0418)	(0.1038)
sdv^j_{t-1}	0.1119**	0.0993***	0.1472***	0.2929***
	(0.0448)	(0.0373)	(0.0322)	(0.0788)
α	0.2596***	0.2356***	0.1665***	0.2902***
	(0.0366)	(0.0360)	(0.0277)	(0.0300)
N	14904	9936	4968	4968
Wald $\chi^2_{(2)}$	58.9	129.8	114.6	65.5

Table B4: Confidence intervals and standard deviation of point forecasts. Note: the table is based on the equation: $sip^k_{t+1|t} = \alpha + \beta sip^k_{t|t-1} + \gamma sdv^j_{t-1} + u^{em}_t$. Coefficients are based on the Blundell-Bond system GMM estimator. Standard errors in parentheses are calculated using bootstrap procedures (1000 replications) that take into account the potential presence of clusters in groups. */**/*** denotes significance at 10/5/1 percent level.

[36] Table B6 in Appendix B reports regression with dummies without interaction with actual forecast errors.

$sip^k_{t+1\|t}$:	*all*
$sip^k_{t\|t-1}$	0.4153***
	(0.0998)
sd^j_{t-1}	0.1034**
	(0.0510)
$T2$	0.9459*
	(0.5606)
$T3$	-0.6684
	(0.6068)
$T4$	-0.6889
	(0.5834)
α	0.3402*
	(0.2976)
N	14904
Wald $\chi^2_{(6)}$	107.5

Table B5: Confidence intervals and standard deviation of inflation Note: coefficients are based on the Blundell-Bond system GMM estimator. Standard errors in parentheses are calculated using bootstrap procedures (1000 replications) that take into account the potential presence of clusters in groups. */**/*** denotes significance at 10/5/1 percent level.

$sip^k_{t+1\|t}$:	*all*	$treat.Sym$	$treat.Asym-L$	$treat.Asym-U$
$sip^k_{t\|t-1}$	0.4636***	0.5719***	0.4780***	0.1090
	(0.1028)	(0.0726)	(0.0532)	(0.1072)
D^k_4	0.0364*	0.0228	0.0054	0.0788**
	(0.0215)	(0.0263)	(0.0121)	(0.0320)
D^k_5	0.0669***	0.0656**	0.0735***	0.0261
	(0.0233)	(0.0295)	(0.0216)	(0.0264)
α	0.2718***	0.2489***	0.1802***	0.3480***
	(0.0376)	(0.0364)	(0.0264)	(0.0400)
N	14688	9792	4896	4896
Wald $\chi^2_{(3)}$	59.0	127.2	138.5	14.5

Table B6: Confidence intervals and the effect of forecast errors. Note: the table is based on the equation: $sip^k_{t+1|t} = \alpha + \beta sip^k_{t|t-1} + \gamma D^k_4 + \delta D^k_5 + u^{em}_t$. Coefficients are based on the Blundell-Bond system GMM estimator. Standard errors in parentheses are calculated using bootstrap procedures (1000 replications) that take into account the potential presence of clusters in groups. */**/*** denotes significance at 10/5/1 percent level.

B.3 Determinants of symmetric confidence intervals

Tables B7 and B8 report results of the following regressions:

$$D_z = \alpha + \beta sip^k_{t|t-1} + \gamma D_1 y_{t-1} + \delta D_2 y_{t-1} + \epsilon D_3 y_{t-1} + \zeta i_{t-1}$$
$$+\eta D_L |\pi_{t-1}| + \theta D_H |\pi_{t-1}| + \phi sd^j_{t-1} + u^{em}_t; \quad z \in \{7, 8, 9\},$$

where $D_7 = 1$ if the $C_U = C_L$ and 0 otherwise, $D_8 = 1$ when $|C_L - C_U| \leq 0.1$, and $D_9 = 1$ when $0.9 \leqslant \left|\frac{ConfIntH_{n-1}}{ConfIntL_{n-1}}\right| \leqslant 1.1$.

logit	D_7	D_8	D_9	D_9, fe
$sip^k_{t\|t-1}$	0.2178	-1.0233*	-0.0194	0.0581
	(0.1957)	(0.6213)	(0.2968)	(0.2239)
$D_1 y_{t-1}$	0.2836	0.4825**	-0.0122	-0.0028
	(0.4162)	(0.2245)	(0.2913)	(0.2846)
$D_2 y_{t-1}$	-0.3912*	0.1116	-0.0490	-0.0605
	(0.2058)	(0.2339)	(0.1847)	(0.1859)
$D_3 y_{t-1}$	-0.3436	0.1057**	-0.0277	-0.0288
	(0.2645)	(0.0441)	(0.1667)	(0.1508)
$D_L\|\pi_{t-1}\|$	0.2375*	0.2203***	0.1635	0.1629
	(0.1354)	(0.0720)	(0.1163)	(0.1143)
$D_H\|\pi_{t-1}\|$	0.1510	0.1858	0.1494	0.1588
	(0.2850)	(0.1214)	(0.1929)	(0.1842)
i_{t-1}	-0.4047	-0.2827**	-0.2041	-0.1969
	(0.2570)	(0.1204)	(0.1660)	(0.1637)
sd^k_{t-1}	-0.1318	-0.4759*	-0.2817**	-0.3295*
	(0.1381)	(0.2477)	(0.1330)	(0.1905)
α	-2.8695***	-0.1098	-1.5233***	-
	(0.4649)	(0.3313)	(0.4104)	
$\ln(\sigma^2_u)$	-0.4665	-1.1088	-0.7481	-
	(0.2516)	(0.2653)	(0.2443)	
σ_u	0.7920	0.5744	0.6879	-
	(0.0996)	(0.0762)	(0.0840)	
ρ^*	0.1601	0.0911	0.1258	-
	(0.0338)	(0.0220)	(0.0269)	
N	4968	4968	4968	4968
Wald $\chi^2_{(8)}$	48.3	72.3	29.2	34.0

Table B7: Determinants of symmetric intervals. Note: coefficients are based on random effects logit estimations, except for "D_9, fe", which is based on fixed effects logit estimation. Standard errors in parentheses are calculated using bootstrap procedures (1000 replications) that take into account the potential presence of clusters in groups. */**/*** denotes significance at 10/5/1 percent level.

Poisson	D_7	D_8	D_9	D_9, fe
$sip^k_{t\|t-1}$	0.1717	-0.7314**	-0.0307	0.0443
	(0.1412)	(0.3667)	(0.2476)	(0.1641)
$D_1 y_{t-1}$	0.2215	0.2141*	-0.0092	0.0032
	(0.3388)	(0.1165)	(0.2046)	(0.1989)
$D_2 y_{t-1}$	-0.2974*	0.0625	-0.0298	-0.0491
	(0.1598)	(0.1137)	(0.1276)	(0.1301)
$D_3 y_{t-1}$	-0.2727	0.0704*	-0.0106	-0.0147
	(0.2216)	(0.0392)	(0.1234)	(0.1050)
$D_L\|\pi_{t-1}\|$	0.1922*	0.1043**	0.1114	0.1106
	(0.1092)	(0.0408)	(0.0820)	(0.0801)
$D_H\|\pi_{t-1}\|$	0.1141	0.0917	0.1004	0.1146
	(0.2330)	(0.0759)	(0.1411)	(0.1353)
i_{t-1}	-0.3303	-0.1336**	-0.1413	-0.1298
	(0.2126)	(0.0619)	(0.1146)	(0.1115)
sd^j_{t-1}	-0.0802	-0.2374***	-0.1933**	-0.2476**
	(0.1356)	(0.0906)	(0.0841)	(0.1203)
α	-2.6585***	-0.6834***	-1.6007***	-
	(0.3787)	(0.1597)	(0.2862)	
$\ln(\alpha^*)$	-0.8804***	-2.8617***	-1.5552***	-
	(0.2106)	(0.9207)	(0.2275)	
α^*	0.4146	0.0572	0.2111	-
	(0.0873)	(0.0526)	(0.0480)	
N	4968	4968	4968	4968
Wald $\chi^2_{(8)}$	40.3	71.3	21.9	27.4

Table B8: Determinants of symmetric intervals. Note: coefficients are based on random effects Poisson estimations, except for "D_9, fe", which is based on fixed effects logit estimation. Standard errors in parentheses are calculated using bootstrap procedures (1000 replications) that take into account the potential presence of clusters in groups. */**/*** denotes significance at 10/5/1 percent level.

B.4 Comparison of different measures: correlation analysis

The aim of this section is to compare different measures of individual uncertainty and disagreement among forecasters. Various studies argue that disagreement measured as the standard deviation of point forecasts lacks a theoretical basis and is therefore not a suitable proxy for uncertainty and consequently also for inflation variability, as is implicit in Zarnowitz and Lambros (1987). However, as we pointed out above, Boero, Smith, and Wallis (2008) question this statement and show that disagreement is a component of the variance of aggregate distribution.

There are several advantages and disadvantages to each measure proposed. The choice of the measure should therefore be oriented to the purpose for which it is intended. Several survey data articles point out that the advantage of the measure of disagreement among forecasters (sdv) is that it is available in any survey, whereas only a limited number of surveys asks for measures of individual uncertainty. Our design thus allows us to also use average confidence interval ($asip$) for comparison. A proxy for the uncertainty may be the average absolute forecast error across individuals (mr). A measure of the variation in the aggregate distribution of forecasts gives information about both uncertainty and disagreement. The interquartile range (IQR) is a proxy for that. Figures A1 and A2 in Appendix A display a timewise comparison between the average confidence interval, the standard deviation of point forecasts and the interquartile range for each group.

We compare pairwise correlation coefficients between different measures of uncertainty and disagreement as in D'Amico and Orphanides (2008) to make a preliminary assessment of their forecasting ability, which is further scrutinized below using dynamic panel regression analysis.

	mr_t^j	$asip_t^j$	sdv_t^j	IQR_t^j	π_{t+1}	i_{t+1}	y_{t+1}
mr_t^j	1						
$asip_t^j$	0.577***	1					
sdv_t^j	0.822***	0.532***	1				
IQR_t^j	0.827***	0.689***	0.777***	1			
π_{t+1}	-0.080**	-0.030	-0.063*	-0.131***	1		
i_{t+1}	0.169***	0.196***	0.226***	0.198***	0.845***	1	
y_{t+1}	-0.321***	-0.246***	-0.259***	-0.267***	-0.016	-0.177***	1
sd_{t+1}^j	0.818***	0.690***	0.722***	0.877***	-0.143***	0.185***	-0.280***

Table B9: Pairwise correlation coefficients. Note: */**/*** denotes significance at 10/5/1 percent level.

All three measures that we compare in this section are significantly positively correlated between each other and with the standard deviation of inflation. However, some of the correlation coefficients are not very high. As we can observe in Table B9, there is a significant correlation coefficient (about 0.5) between the average width of the confidence interval and the standard deviation of point forecasts.[37] Rich and Tracy (2010) and Boero, Smith, and Wallis (2008) find little evidence that this relationship exists in the survey data, while D'Amico and Orphanides (2008) find a correlation coefficient of 0.4. The present analysis suggests that uncertainty and disagreement are modestly correlated.

A positive correlation between the interquartile range and individual uncertainty can be observed. The correlation coefficient (around 0.7) is higher than that reported in the previous paragraph. As shown in the statistical analysis by Boero, Smith, and Wallis (2008), there exists a "structural" relationship between these two variables so a positive relationship is expected. For similar reasons there is also a correlation between the disagreement and the interquartile range. The latter correlation is of similar magnitude to the former. Therefore, one could argue that the interquartile range is in our experiment at least as much, if not more, a measure of disagreement as average individual uncertainty. Bomberger (1996) argues that the standard deviation of point forecasts is a useful proxy for uncertainty and that disagreement tracks uncertainty better than the GARCH model; however, this view is questioned by Rich and Butler (1998).[38]

[37] Table B4 in Appendix B depicts the relationship between confidence bounds and the dispersion of point forecasts in more detail. We find no evidence of this relationship for symmetric intervals, while for asymmetric there is a positive relationship.

[38] Lahiri and Sheng (2010) point out that disagreement is useful for forecasting in stable periods but not in periods of high volatility.

C Confidence intervals under Rational Expectations

To calculate the standard deviation of inflation forecast uncertainty under Rational Expectations (RE), we have to first define the perceived law of motion (PLM) of the RE equilibrium form (under the assumption of homogeneous expectations):

$$E_t\pi_{t+1} = R_0 + R_1 y_{t-1} + R_2 g_{t-1} + R_3 u_{t-1}, \tag{15}$$

where all coefficients R are calculated using the method of undetermined coefficients. See our companion paper for details (Pfajfar and Žakelj, 2011). Note that $E_t\pi_{t+1}$ are inflation forecast made in t with information from $t-1$ for period $t+1$.[39] To calculate inflation forecast uncertainty we have to feed forward for one period the equations (1), (2), (15) and the relevant monetary policy rule (either (3) or (4) depending for which regime we are calculating it). Then we insert (15), (1) and policy rule into the forwarded (1) and the resulting equation into the (2). We thus obtain the actual law of motion (ALM) for inflation. The unobservable part of the shock of this ALM represents the inflation forecast uncertainty. Thus, the standard deviation of forecast uncertainty for policy rules $p = 1, ..., 3$ is:

$$\begin{aligned}
\sigma_k &= \sqrt{\left((C(R_1+R_2)+\lambda(1+\kappa))^2+\lambda^2\right)\sigma_g^2+\left((C\cdot R_3+\nu)^2+1\right)\sigma_u^2}, \\
C &= \lambda\varphi(1-\gamma)+\beta.
\end{aligned}$$

where σ_k is the standard deviation of inflation forecast uncertainty of the forecaster k, who uses rational expectations to forecast inflation. Standard deviation of forecast

[39] In this and the next section we denote $E_t\pi_{t+1}$ made by a forecaster k as π^k.

uncertainty for policy rules $p = 4$ is:

$$\begin{aligned}
\sigma_k &= \sqrt{(C_2^2 + C_1^2)\,\sigma_g^2 + (C_3^2 + (1 - C_1\varphi\gamma)^2)\,\sigma_u^2}, \\
C_1 &= \lambda/(1 + \varphi\gamma\lambda), \\
C_2 &= C_1\kappa + C_1^2/\lambda + (C_1\varphi(1 - \gamma\beta) + \beta)\,(R_2 + C_1 R_1/\lambda)\,, \\
C_3 &= \nu(1 - C_1\varphi\gamma) - C_1^2\varphi\gamma/\lambda + (C_1\varphi(1 - \gamma\beta) + \beta)\,(R_3 - C_1\varphi\gamma R_1/\lambda)\,.
\end{aligned}$$

Numerical values are reported in Table D1.

D Properties of the payoff function

In this Appendix we study the properties of the payoff function W_2. Profit, conditional on the actual value of inflation, π, falling within confidence interval CI from the forecasted value, π^k is:

$$W_2 = \frac{100}{1 + CI} - 20.$$

Probability that the actual value π falls within a confidence interval CI from the forecasted value, π^k :

$$P_W = F(CI) - F(-CI) = \int_{-CI}^{CI} f\,(x)\,dx,$$

where $f\,(x)$ is a probability density function of a forecasting error, $x = \pi^k - \pi$, and $F(x)$ a corresponding cumulative density function. Expected profits, EW, are therefore:

$$\begin{aligned}
EW_2 &= W_2 \cdot P_W, \\
EW_2 &= \left(\frac{100}{1 + CI} - 20\right)(F(CI) - F(-CI))\,.
\end{aligned}$$

Maximization condition is thus:

$$\frac{\partial EW_2}{\partial d} = W_2' \cdot P_W + W_2 \cdot P_W' = 0.$$

For the rational expectations and the restricted perception equilibrium forecast errors are distributed normally, $f(\pi_k - \pi) \sim N(0, \sigma_k^2)$. σ_k^2 is computed in Appendix C and is reported in Table D1 for rational expectations and in Table D2 for restricted perception equilibrium. Derivatives are:

$$
\begin{aligned}
W_2 &= \frac{80 - 20CI}{1 + CI} \\
W_2' &= -\frac{100}{(1 + CI)^2} \\
P_W &= \frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{CI}{\sigma_k\sqrt{2}}\right) - 1 - \operatorname{erf}\left(\frac{-CI}{\sigma_k\sqrt{2}}\right)\right] \\
&= \frac{1}{2}\left[\operatorname{erf}\left(\frac{CI}{\sigma_k\sqrt{2}}\right) + \operatorname{erf}\left(\frac{CI}{\sigma_k\sqrt{2}}\right)\right] = \operatorname{erf}\left(\frac{CI}{\sigma_k\sqrt{2}}\right) \\
P_W' &= \sqrt{\frac{2}{\pi\sigma_k^2}}e^{-\frac{CI^2}{2\sigma_k^2}} = \frac{2}{\sigma_k^2\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{CI}{\sigma_k}\right)^2} \\
&= 2 \cdot f(CI; 0, \sigma_k^2)
\end{aligned}
$$

Thus,

$$
\frac{\partial EW_2}{\partial d} = -\frac{100}{(1 + CI^*)^2}\operatorname{erf}\left(\frac{CI^*}{\sigma_k\sqrt{2}}\right) + \frac{80 - 20CI^*}{1 + CI^*}\frac{2}{\sigma_k^2\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{CI^*}{\sigma_k}\right)^2} = 0
$$

Solving this equation for confidence intervals, CI^* under RE for different treatments gives us the following values:

Taylor rule (equation)	Parameters	σ_k	CI^*	$ICMP^*$
1 – Forward looking (3)	$\gamma = 1.5$	0.3511	0.5050	85%
2 – Forward looking (3)	$\gamma = 1.35$	0.4161	0.5633	82%
3 – Forward looking (3)	$\gamma = 4$	0.1647	0.2958	93%
4 – Contemporaneous (4)	$\gamma = 1.5$	0.2740	0.4272	88%

Table D1: Incentive compatibility of the payoff function under Rational Expectations

From P_W we can also deduct the percentage of cases where the actual inflation fall within the confidence bounds when this function is maximized $(ICMP^*)$. We can also study incentive compatibility under restrictive perception equilibrium, where agents, as in our experiment do not observe the shocks (but is otherwise of the same form as rational expectations equilibrium). Under this expectation formation mechanism the results are the following:

Taylor rule (equation)	Parameters	σ_k	CI^*	$ICMP^*$
1 – Forward looking (3)	$\gamma = 1.5$	0.2273	0.3768	90%
2 – Forward looking (3)	$\gamma = 1.35$	0.2587	0.4175	89%
3 – Forward looking (3)	$\gamma = 4$	0.1397	0.2575	94%
4 – Contemporaneous (4)	$\gamma = 1.5$	0.1953	0.3350	91%

Table D2: Incentive compatibility of the payoff function under Restrictive Perception Equlibria

E Instructions for the experiment (not for publication)

Thank you for participating in this experiment, a project of economic investigation. Your earnings depend on your decisions and the decisions of the other participants. There is a show up fee of 5 Euros assured. From now on until the end of the experiment you are not allowed to communicate with each other. If you have any questions raise your hand and one of the instructors will answer the question in private. Please do not ask aloud.

The experiment

All participants receive exactly the same instructions. You and 8 other subjects all participate as agents in the *same* fictitious economy. You will have to predict future values of given economic variables. The experiment consists of 70 periods. The rules are the same in all the periods. You will interact with the same 8 subjects during the whole experiment.

Imagine that you work in a firm where you have to predict inflation for the next period. Your earnings depend on the accuracy of your inflation expectation.

Information in each period

The economy will be described with 3 variables in this experiment: the *inflation rate*, the *output gap*, and the *interest rate*.

- **Inflation** measures a general rise in prices in the economy. In each period it depends on the inflation expectations of the agents in economy (you and the other 8 participants in this experiment), the output gap and random shocks which have equal probability of having a positive or negative effect on inflation and are normally distributed.

- The **output gap** measures by how much (in percentage) the actual Gross Domestic Product differs from the potential one. If the output gap is greater than 0, it means that the economy is producing more than the potential level, if negative, less than the potential level. In each period it depends on the inflation expectations of the agents in the economy, the past output gap, the interest rate and random shocks which have equal probability of having a positive or negative effect on inflation and are normally distributed.

- The **interest rate** is (in this experiment) the price of borrowing money (in percentage) for one period. The interest rate is set by the monetary authority. Their decision mostly depends on the inflation expectations of the agents in the economy.

All given variables might be relevant to your inflation forecast, but it is up to you to work out their relation and the possible benefit of knowing them. The evolution of the variables will partly depend on your and the other subjects' inputs and also different exogenous shocks influencing the economy.

- You enter the economy in period 1. In this period you will be given computer-generated past values of inflation, the output gap and the interest rate for 10 periods back (Called: -9, -8, ... -1, 0)

- In period 2 you will be given all the past values as seen in period 1 plus the value from period 1 (Periods: -9, -8, ... 0, 1).

- In period 3 you will see all the past values as in period 2 (Periods: -9, -8, ... 1, 2) plus YOUR prediction about inflation in period 2 that you made in period 1.

- In period t you will see all the past values of actual inflation up to period $t-1$ (Periods: -9, -8, ... $t-2$, $t-1$) and your predictions up to period $t-1$ (Periods: 2, 3, ... $t-2$, $t-1$).

What do you have to decide?

Your task is to predict the state of the economy as accurately as possible. Your payoff will depend on the accuracy of your prediction of the inflation in the future period. In each period your prediction will consist of two parts:

a) *Expected inflation*, (in percentage) that you expect to be in the NEXT period ($Exp.Inf.$)

b) *Lower bound* (in percentage) of your prediction. You must be almost sure that the actual inflation will be higher than your lower bound.

c) *Upper bound* (in percentage) of your prediction. You must be almost sure that the actual inflation will be lower than your upper bound.

Based on b) and c) we determine the confidence interval, $Conf.Int.$ which is equal to

$$Conf.Int. = Upper\ bound\ -\ Lower\ bound$$

Example 1 *Let's say you think that inflation in the next period will be* 3.7%. *And you also think it is most likely (*95% *probability) that the actual inflation will not be lower than* 3.2% *and not higher than* 4.0%. *Your inputs in the experiment will be* 3.7 *under a),* 3.2 *under b), and* 4.0 *under c).*

Your goal is to maximize your payoff, given with the equation:

$$W = \max\left\{\frac{100}{1+|Inflation - Exp.Inf.|} - 20, 0\right\} + \max\left\{\frac{100x}{1+\frac{1}{2}Conf.Int.} - 20, 0\right\}$$

where $Exp.Inf.$ is your expectation about the inflation in the NEXT period, $Conf.Int.$ is the confidence interval, Inflation is the actual inflation in the next period and x is a variable with value 1 if

$$Lower\ bound \leq Inflation \leq Upper\ bound$$

and 0 otherwise.

The *first part* of the payoff function states that you will receive some payoff if the actual value in the next period differs from your prediction in this period by less than 4 percentage points. The smaller this difference is, the higher the payoff you receive. With a zero forecast error ($|Inflation - Exp.Inf.| = 0$), you would receive 80 units ($100/1 - 20$). However, if your forecast is 1 percentage point higher or lower than the actual inflation rate, you will get only 30 units ($100/2 - 20$). If your forecast error is 4 percentage points or more, you will receive 0 units ($100/5 - 20$).

The *second part* of the payoff function simply states that you will get some extra payoff if the actual inflation is within your expected interval and if that interval is not larger than 8 percentage points. The more certain of the actual value you are, the smaller interval you give (*Lower bound* and *Upper bound* closer to $Exp.Inf.$), and the higher will be your payoff if the actual inflation is indeed in the given interval, but there will also be higher chances that the actual value will fall outside your interval. In our example this interval was 0.8 percentage points. If the actual inflation falls in this interval you receive 51.4 units ($100/(1 + \frac{1}{2}0.8) - 20$) in addition to the payoff from the first part of the payoff function. If the actual values is outside your interval, your receive 0.

On the attached sheet you will find a table showing various combinations of *forecast error* and *confidence interval* needed to earn a given number of points.

Information after each period

Your payoff depends on your predictions for the next period and the actual realization in the next period. Because the actual inflation will be known only in the next period, you will also be informed about you current period (t) prediction and earnings after the end of the NEXT period ($t + 1$). Therefore:

- After period 1 you will not receive any earnings, since you did not make any prediction for period 1.
- In any other period, you will receive information about the actual inflation rate in this period and your *inflation* and *confidence interval* prediction from the previous period. You will also be informed if the actual inflation value was in your expected interval and what your earnings are for this period.

The units in the experiment are fictitious. Your actual payoff (in euros) will be the sum of earnings from all periods divided by 500.

If you have any questions please ask them now!

Questionnaire[40]

1. If you believe that inflation in the next period will be _ _4.2%_ _, and you are quite sure that it will not go down by more than _ _ 0.4_ _ nor up by more than _ _ 0.7_ _, you will type:
   Under (1) _ _ _ _ _ _ _ _ _ _ for inflation,
   Under (2) _ _ _ _ _ _ _ _ _ _ for the lower bound, and
   Under (3) _ _ _ _ _ _ _ _ _ _ for the upper bound.
2. You are now in period _ _ _ _15_ _ _ _. You have information about past inflation, the output gap and the interest rate up to period _ _ _ _ _ _ _ _ _ _ and you have to predict the inflation for period _ _ _ _ _ _ _ _ _ _.

[40] Options (1), (2) and (3) point to the different fields on the screenshot of the experimental interface.